

Modelling, Monitoring and Validation of Plant Phenology Products

Von der Fakultät für Physik und Geowissenschaften
der Universität Leipzig
genehmigte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt

von M.Sc. Maximilian Lange

geboren am 24.09.1988 in Halle/Saale

Gutachter: Prof. Dr. Hannes Feilhauer

Prof. Dr. Michael Vohland

Tag der Verleihung: 16. Dezember 2019

Bibliographische Beschreibung:

Lange, Maximilian

Modelling, Monitoring and Validation of Plant Phenology Products

Universität Leipzig, Dissertation

216 S., 261 Lit., 45 Abb., 17 Tab., 2 Anlagen

Referat Phänologie, die Lehre der periodisch wiederkehrenden Entwicklungserscheinungen in der Natur, hat sich in den letzten Jahrzehnten zu einem wichtigen Teilgebiet der Klimaforschung entwickelt. Einer der Haupteffekte der globalen Erwärmung ist die Veränderung der Wachstumsmuster und Fortpflanzungsgewohnheiten von Pflanzen, und somit veränderte Phänologie. Um die Auswirkungen der Klimaveränderung auf wildwachsende sowie Kulturpflanzen vorherzusagen, werden phänologische Modelle angewendet, verbessert und validiert. Dabei ist Wissen über den aktuellen Stand der Vegetation notwendig, welches aus Beobachtungen und fernerkundliche Messungen gewonnen wird. Die hier präsentierte Arbeit befasst sich mit dem Verständnis der Zusammenhänge zwischen fernerkundlichen Messungen und phänologischen Stadien und somit den Herausforderungen der modernen phänologischen Forschung: Der Vorhersage der Phänologie durch Modellierungsansätze, der Beobachtung der Phänologie mit optischen boden- und satellitengestützten Sensoren und der Validierung phänologischer Produkte.

Abstract Phenology, the study of recurring life cycle events of plants and animals has emerged as an important part of climate change research within the last decades. One of the main effects of global warming on vegetation is altered phenology, since plants have to modify their growth patterns and reproduction habits as reaction to changing environmental conditions. Forecasting phenology, thus phenological modelling, is a timely challenge given the necessity to predict the impact of global warming on wild-growing species and agricultural crops. However, assessing the present state of vegetation, thus phenological monitoring, is essential to update and validate model results. An improved comprehension of the relationships between plant phenology and remotely sensed products is crucial to interpret these results. Consequently, the presented thesis deals with the main challenges faced in modern phenology research, covering phenological forecasting with a modelling approach, satellite-based phenology extraction, and near-surface long-term monitoring of phenology.

Acknowledgements

Firstly, I would like to thank my thesis advisor Prof. Dr. Michael Vohland of the Institute of Geography, Geoinformatics and Remote Sensing at Leipzig University. The door to the office of Prof. Vohland was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

I would like to express my sincere gratitude to my advisor Dr. Daniel Doktor of the Department Computational Landscape Ecology at the Helmholtz-Centre for Environmental Research - UFZ for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study.

My sincere thanks also goes to Dr. Franz-Werner Badeck, Dr. Matthias Cuntz, Greta Jäckel, Dr. Dirk Koslowsky, Dr. Andreas Marx, Dr. Corinna Rebmann, Dr. Jörg Schaber and Prof. Dr. Ralf Seppelt for their insightful comments and encouragement, but also for the hard questions which incited me to widen my research from various perspectives. Without their precious support it would not be possible to conduct this research.

I would like to express my gratitude to the technicians of the Departments Computational Landscape Ecology and Computational Hydrosystems at UFZ, especially Gudrun Schuhmann, Gundula Schulz, Steffen Lehmann, Sebastian Gimper and Hendrik Zöphel for their technical support and precious work on the infrastructure this thesis was based on. In addition I would like to thank the staff of the scientific and commercial data processing unit at UFZ, especially Ben Langenberg and Christian Krause, for their support and immense knowledge. Without them, scientific computing would be a closed book for me. I am also grateful to my fellow doctoral students and colleagues Benjamin Dechant, Niklas Hase, Sebastian Preidl and Andreas Schmidt for their feedback, cooperation and of course friendship.

Special thanks should be given to my friends. They incited me to strive towards my goal and have been pretty understanding, patient and encouraging throughout the process. Last but not the least, I would like to thank my family: my parents and my girlfriend Andrea for their continuous encouragement throughout my years of study and through the process of writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

1	Introduction to Plant Phenology	1
1.1	Phenological Modelling	3
1.2	Satellite Phenology Monitoring	5
1.3	Ground-based Phenological Measurements	8
1.4	Study Area	9
1.5	Research Objectives and Hypothesis	12
2	Simulation of Forest Tree Species' Bud Burst Dates for Different Climate Scenarios: Chilling Requirements and Photo-Period May Limit Bud Burst Advancement	17
	Acknowledgements	18
	Author Contributions	18
	Abstract	18
2.1	Introduction	19
2.2	Data and Methods	20
2.2.1	Temperature Data	20
2.2.2	Phenological Data	21
2.2.3	Promoter-Inhibitor-Model	23
2.2.4	Fitting with DEoptim	27
2.2.5	Sensitivity Analysis	28
2.3	Results	29
2.3.1	Combined Phenological Time Series	29
2.3.2	Sensitivity Analysis and Model Fits	30
2.3.3	Model Performance	32
2.3.4	Climate Scenarios	32

2.4	Discussion	36
2.5	Conclusions	40
	Supporting Information	40
Sup.1	Selected Promoter-Inhibitor-Models	40
Sup.2	DEoptim Parameter Settings	41
3	Mixed pixel effects and land-cover type translate method dependent into varying spring phenology trend estimates using coarse resolution satellites	45
	Acknowledgements	46
	Author Contributions	46
	Abstract	46
3.1	Introduction	47
3.2	Materials and Methods	51
3.2.1	Satellite data	51
3.2.2	Ground observational data	52
3.2.3	Bud burst modelling	54
3.2.4	Simulated NDVI time series	54
3.2.5	Extracting phenological metrics	57
3.2.6	Trend analysis	59
3.2.7	Goodness-of-fit	59
3.3	Results	60
3.3.1	Satellite vs ground observations	60
3.3.2	Trends from simulated NDVI profiles	60
3.3.3	Goodness-of-fit	61
3.3.4	Green-up dates across methods	62
3.3.5	Trends from satellite observations	66
3.4	Discussion and conclusions	71
3.4.1	Trends from simulated NDVI time series	71
3.4.2	Satellite data trend estimates	72
3.4.3	Comparison to studies using method ensembles	73
3.4.4	The satellite's spatio-temporal resolution	74
3.4.5	Ground truth & modelling	75
3.4.6	Sensor inter-calibration and potential false trends	75

3.4.7 Recommendations	77
Data accessibility	78
4 Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors	79
Acknowledgements	80
Author Contributions	80
Abstract	80
4.1 Introduction	81
4.2 Materials and Methods	84
4.2.1 Study Site and Sensor Setup	84
4.2.2 Generation of NDVI Products from Ground-Based Spectral Measurements	86
4.2.3 Satellite Data and Respective NDVI Products	88
4.2.4 NDVI Post-Processing and Phenological Metrics Extraction	89
4.3 Results	91
4.3.1 The NDVI Products at Different Scales	91
4.3.2 Analysis of Phenological Metrics	94
4.4 Discussion	96
4.5 Conclusions	100
5 Continuous calibration monitoring of automated hyperspectral measurements of vegetation canopy optical properties. I. Radiometric and reflectance calibration	103
Acknowledgements	104
Author Contributions	104
Abstract	104
5.1 Introduction	105
5.2 Materials and Methods	108
5.2.1 Study sites	108
5.2.2 Measurement set-up	108
5.2.3 Calibration, monitoring and correction of hyperspectral systems	110
5.3 Results	114
5.3.1 Campaign-based white reference measurements for calibration stability monitoring at the grassland site	114

5.3.2	Continuous stability monitoring of hyperspectral observations . . .	115
5.3.3	Continuous stability correction of hyperspectral observations	118
5.3.4	Comparison of hyperspectral reflectance observations with calibrated field spectrometer and example spectra during phase of rapid phenological changes	120
5.4	Discussion	121
5.4.1	Potential and likely causes of calibration instability of hyperspectral systems	121
5.4.2	Spectral stability of reflectance correction factor	123
5.4.3	Stability of multispectral measurements	124
5.4.4	Performance of the continuous correction of hyperspectral observations	125
5.4.5	Comparison of different hyperspectral system set-ups with respect to reflectance calibration stability	127
5.4.6	Limitations of our study	128
5.4.7	Implications of our results and different correction methods	129
5.5	Conclusions	130
	Supplemental Material	132
6	Discussion	139
6.1	Modelling bud burst in a future warming climate	140
6.2	Comparison of methods extracting phenological metrics from satellite data .	141
6.3	Validation of phenology products with near-surface sensor-derived data . .	143
6.4	Continuous calibration of unattended near-surface sensor-systems	145
6.5	Synopsis of modelling, monitoring and validation of plant phenology products	146
6.6	Conclusions	148
	Bibliography	151
	List of Figures	183
	List of Tables	189
	List of Abbreviations and Variables	193
	Declaration of Authorship	203

1 Introduction to Plant Phenology

Phenology is the study of periodically recurring growth phenomena of plants and animals and the timing of their occurrence (Schnelle, 1955). The name is derived from the Greek words *phainō* ("to appear") and *logos* ("science"). For plants, recurring phenomena are mainly development stages, such as bud burst, leaf unfolding, blossoming, fruit ripening and leaf colouring (Doktor, 2008). The oldest known time series of phenology worldwide are records of cherry blossom in Japan, starting in 705 AD (Lauscher, 1978). These records served the timing of the cherry blossom festival rather than science. In Europe, the oldest known records are dates of grape harvesting in 1484 (Menzel, 2005). Modern systematic phenology started with the description of purpose and method of phenological observations by Carl von Linné in 1751, followed by the establishment of the first regional phenological network (Doktor, 2008). The first international phenological network, established by the Societas Meteorologica Palatina in Mannheim, Germany, was active between 1781 and 1792 (Schnelle, 1955). Hoffmann and Ihne established a broad European phenological observation network in 1882, including precise observation instructions and recordings on a regular basis (Schaber, 2002). Observations were recorded until 1941. A volunteer network started systematic observation of phenology in Germany in 1922. It was taken over by the Deutscher Reichswetterdienst in 1936 and continued by the German Weather Service (Deutscher Wetterdienst, DWD) and Hydrometeorologischer Dienst der DDR after the Second World War until 1991. DWD and Hydrometeorologischer Dienst der DDR were unified in 1991 and managed by the DWD until today (Menzel, 2013a).

Phenological networks usually served agrometeorology by selecting regions best suited for cultivation of certain crops and fruits, as well as by coordinating the use of harvesting machines (Doktor, 2008). The knowledge gained from phenology was used for plant protection,

late frost risk reduction, timing of herbicide and pesticide application, as well as horticulture and viticulture (Schaber, 2002; Chmielewski, 2013). At the end of the 20th century, researchers began to use phenological observations to study climate change (Doktor, 2008; Schwartz, 2013d). Many studies observed rising global surface temperatures (Hansen et al., 1999, 2006, 2010). The IPCC (Intergovernmental Panel on Climate Change) concluded with high confidence that humans contributed substantially to global warming since at least 1750 (Intergovernmental Panel on Climate Change, 2007; Fussel et al., 2012). One of the main effects of global warming on vegetation is altered phenology, since plants have to modify their growth patterns and reproduction habits as reaction on changing environmental conditions (Kramer et al., 1996; Parmesan, 2007; Sherry et al., 2007). These phenological shifts have already been detected in ground observational (Fussel et al., 2012; Intergovernmental Panel on Climate Change, 2014) and remote sensing data (Barichivich et al., 2013; Fu et al., 2014). Potential impacts of phenological shifts are numerous (Intergovernmental Panel on Climate Change, 2014), mainly including an enhanced vulnerability of plants to cold spells (Schaber and Badeck, 2003; Linkosalo et al., 2006) and late frost (Kramer et al., 1996; Fussel et al., 2012), consequences for species composition and distribution (Thackeray et al., 2010; Peñuelas et al., 2013) and a subsequently altered carbon exchange of ecosystems (Peñuelas et al., 2013). Thus, modelling phenology, briefly introduced in Section 1.1, improves insight into long-term consequences of changing environmental conditions (Peñuelas et al., 2009; Schwartz, 2013b). On the other hand, monitoring phenology is a timely challenge given the necessity to calibrate and validate models, as well as to assess the current impact of climatic conditions.

Monitoring phenology includes the common approach of human observations as well as the modern approach of remote sensing, using optical sensors to gather information contained in light reflected from a plants surface in order to infer the vegetation state (Doktor, 2008; Schwartz, 2013c). Human phenology observations often show high uncertainties, mainly stemming from so called observer errors, including varying observation instructions, subjective interpretation and logging errors, as well as genetic differences between the observed plants and the large natural variability of phenological events, which may sum up to uncertainties of several weeks (Schaber, 2002). Further, they describe the onset of phenological phases at the level of individual plants, whereas remotely sensed phenological records represent a spatially integrated response of vegetation within the optical sensors' field of view (Doktor, 2008). Hence, human observations may be scaled to a level of varieties, provenances or species, rather than landscapes with mixtures of species (Doktor, 2008). Remote

sensing provides the possibility to analyse phenology and its relationships with climate on a much larger scale. A brief overview about satellite and ground level phenology observations is given in Section 1.2 and 1.3, respectively. However, remote sensing derived phenology metrics lack universal definitions and consequently differ between products, beside quantifiable measurement uncertainties. Thus, human observations are still highly valuable and essential for the validation of remote sensing products as well as phenological models.

1.1 Phenological Modelling

Modelling phenology dates back to the work of Réaumur (1735), but was not widely used before the availability of computers (Hänninen and Kramer, 2007). These early phenological models, e.g. the temperature sum (day degree) model (Arnold, 1959; Wang, 1960), related the rate of development to one or more environmental conditions. Since the 1970's, more comprehensive models added the prevailing state of development as an influencing factor, enabling the inclusion of rest phenomena in phenological models (Hänninen and Kramer, 2007). During the last three decades, phenological modelling was increasingly used in human and public health for predicting mass proliferation of pollen or vector-borne diseases like malaria (Schaber, 2002), in agriculture for food production and plant protection (Doktor, 2008), and in climate change research (Hänninen and Kramer, 2007).

Early approaches were limited by not being applicable to both the most important species and wide geographic regions, mainly due to the availability of phenological and meteorological data, as well as processing power (Doktor, 2008; Hänninen and Kramer, 2007). Further, physiological processes within plants controlling their phenological development were largely unknown (Schaber, 2002; Chuine et al., 2013). Modern approaches are now able to use extensive phenological databases (Dierenbach et al., 2013; Chmielewski et al., 2013), meteorological information is widely available and processing power increases steadily. These improved conditions led to the formulation of new mechanistic models including physiological processes (Hänninen and Kramer, 2007; Fu et al., 2012), e.g. the sequential (Hänninen, 1990; Kramer, 1994), parallel (Landsberg, 1974; Hänninen, 1990), alternating (Murray et al., 1989) and unified model (Chuine, 2000), as well as the promoter-inhibitor-model (PIM, Schaber and Badeck (2003)). The latter focusses on species specific parameters in order to being applicable for a wide range of conditions. Hence, the model is not limited to specific regions or environmental conditions to which it was fitted, enabling a depiction of

phenological behaviour in future climates (Lange et al., 2016).

Two methodological approaches are used to estimate model parameter values of mechanistic models (Chuine et al., 2013): the experimental approach, analysing the underlying mechanisms of phenological response, and the numerical approach, estimating model parameters through statistical model-fitting techniques relating meteorological data to the timing of the observed phenological events. The phenological data used for development and testing of models originate from two main sources: observations of wild populations or from phenological gardens and experiments (Chmielewski et al., 2013; Chuine et al., 2013). Meteorological information is either obtained from weather stations in the vicinity of the observation or measured directly at the experimental site, respectively (Chmielewski et al., 2013; Chuine et al., 2013).

Most present large scale phenological models for wild-growing species in temperate climates base on ambient temperature and day length, respectively photo-period, which influence growth behaviour of plants significantly (Schnelle, 1955). The role of photo-period in plant phenology is ambiguous (Chuine et al., 2010), although there is evidence that long photo-period compensates for a lack of chilling during dormancy (Hänninen and Tanino, 2011). It has been found that winter precipitation could play a role for spring phenology (Fu et al., 2014). However, due to the lack of data and the requirement to include interactions with soil, water availability is not included in most phenological models. Most differences in phenological models stem from varying response functions to temperature and photo-period (Chuine et al., 2013) and from different considered development phases (Chuine et al., 2013), including dormancy, active vegetation, germination (which depends also on soil moisture), flowering and maturation. Plant phenological models mainly target at depicting bud burst or flowering, since the timing of these phases can be observed accurately (Chuine et al., 2013) and is considered as an indicator of climate change (Intergovernmental Panel on Climate Change, 2014). An increasing number of studies reports an earlier onset of greening (Vitasse et al., 2011; Migliavacca et al., 2012; Intergovernmental Panel on Climate Change, 2014) with a stagnating advancement in the 21th century considering the predicted spring warming (Morin et al., 2009; Fu et al., 2015b). Recent studies reveal that there is a considerable inter-specific diversity in temperature response (Morin et al., 2009; Chuine et al., 2010; Hänninen and Tanino, 2011; Menzel, 2013b). Biological realism may be necessary in order to improve the mechanistic comprehension of plant phenology (Hänninen and Kramer, 2007) and especially the complex interactions between chilling requirements, temperature and photo-period (Lange et al., 2016).

1.2 Satellite Phenology Monitoring

In the late 1970's, the National Aeronautics and Space Administration (NASA) launched the weather satellite TIROS-N containing the first Advanced Very High Resolution Radiometer (AVHRR, NOAA (2017)). The US National Oceanic and Atmospheric Administration (NOAA) managed the system in order to remotely sense cloud cover and surface temperature (NOAA, 2017). AVHRR was originally not designed to study vegetation and its potential for this purpose was unexpected (Cracknell, 1997), but soon recognized. NOAA started the production of a global vegetation index (GVI) in 1982 (Tarpley, 1991). The surprising success of AVHRR for vegetation studies bases on reflection characteristics of healthy plants: light is reflected in the green (around 550 nm) and near-infrared (NIR, >700 nm) part, whereas it is absorbed in the blue (around 430nm) and red (around 650nm) part of the spectrum (Gitelson et al., 2003; Doktor, 2008). These characteristics are different for damaged or dying vegetation (Doktor, 2008). Thus, the magnitude of reflection in this part of the electromagnetic spectrum can be used to identify the vegetation condition (Doktor et al., 2009). The Normalized Difference Vegetation Index (NDVI) is a part of the GVI (Tarpley, 1991) and was developed to exploit these characteristics (Tucker, 1979; Doktor et al., 2009). The seasonal progression of the NDVI of deciduous vegetation in temperate climates is characterized by a rapid increase in spring during leaf development and maturation, a slight decrease during summer, a rapid decrease in autumn due to leaf fall and a low value in winter due to vegetation inactivity (Doktor et al., 2009). AVHRR is able to capture these seasonal characteristics due to its fine temporal resolution. However, due to its coarse spatial resolution landscape details are blurred and canopies can not be resolved as scene objects (Henebry and de Beurs, 2013). Thus, the term land surface phenology (LSP) was introduced, describing the seasonality of reflectance characteristics of vegetated landscapes (White and Nemani, 2006; Henebry and de Beurs, 2013).

Several studies used the GVI product and the NDVI within to investigate different methods deriving LSP (Reed et al., 1994; Lüdeke et al., 1996; Botta et al., 2000) or monitoring plant growth (Myneni et al., 1997). Other vegetation indices, such as the EVI (Enhanced Vegetation Index) or PPI (Plant Phenology Index), were also used to derive LSP (Jin and Eklundh, 2014). Although the NDVI exhibits a sensitivity to snow and may saturate in dense forest canopies, its robustness against varying illumination geometries and noise made it one of the most widely-used indices for phenology depiction (Doktor, 2008; Jin and Eklundh, 2014). Subsequently, new satellite generations and missions were launched in order to intentionally

observe vegetation state and surface phenology, including new generation AVHRR, LANDSAT (USGS, 2018), SPOT (Maisongrande et al., 2004), NASA's Aqua and Terra satellites carrying the Moderate Resolution Imaging Spectroradiometer (MODIS, NASA (2017)) and the European Space Agency's (ESA) Sentinel-2 (Fletcher, 2012), among others. AVHRR, MODIS and Sentinel-2 data are part of this thesis and further introduced.

The second generation AVHRR was launched with NOAA satellites since 1981, incorporating five spectral bands (Kidwell, 1995; NOAA, 2017). The third generation AVHRR with six bands was launched with NOAA satellites since 1998 (Robel and Graumann, 2014; NOAA, 2017). In 1999, the NASA launched its Terra satellite, carrying MODIS. The spatial resolution for vegetation studies was enhanced: whereas AVHRR showed a spatial resolution of 1.1km at nadir (Kidwell, 1995; Robel and Graumann, 2014; NOAA, 2017), MODIS records data with up to 250m resolution in two bands and 500m in five of its 36 bands (NASA, 2017). In comparison to AVHRR, sensor calibration and atmospheric correction were improved, mainly due to the incorporation of spectral bands important for atmospheric content estimation, as well as geolocation accuracy (Soudani et al., 2008; NASA, 2017). The second NASA satellite Aqua carrying MODIS was launched in 2002 (NASA, 2017). In 2015 and 2017, the ESA launched its Sentinel-2 satellites A and B, respectively, carrying the MultiSpectral Instrument (MSI) with 13 bands (Fletcher, 2012). The MSI shows a spatial resolution of 10m and 20m for bands important for vegetation studies and 60m for its aerosol band (Fletcher, 2012), further improving atmospheric correction and the applicability for vegetation studies (see Figure 1.1).

In recent studies, LSP was used for climate change research, since it impacts carbon, water and nitrogen cycles (White and Nemani, 2006; White et al., 2009; Gonsamo et al., 2012). Methodologies to estimate LSP are numerous, but without a universal definition of underlying metrics: green-up, leaf-out and start of season or start of spring (SOS) appear to describe the same event, but may represent different processes (White et al., 2009). In addition, considerable differences in results are reported when extracting the SOS from annual time series, mainly due to the choice of algorithms (Fu et al., 2014). Thus, inter-comparison of studies analysing trends in phenological metrics is complicated. Further problems arise due to the inclusion of atmospheric contamination, cloud cover, snow cover, soil wetness and bi-directional viewing effects in satellite data (White et al., 2009). There is consensus in LSP studies analysing NDVI data to restore the upper envelope of the NDVI curve in order to remove noise, since clouds and other non-vegetational effects usually decrease NDVI values (Doktor, 2008). Maximum Value Composite (MVC, Holben (1986)) is a common method

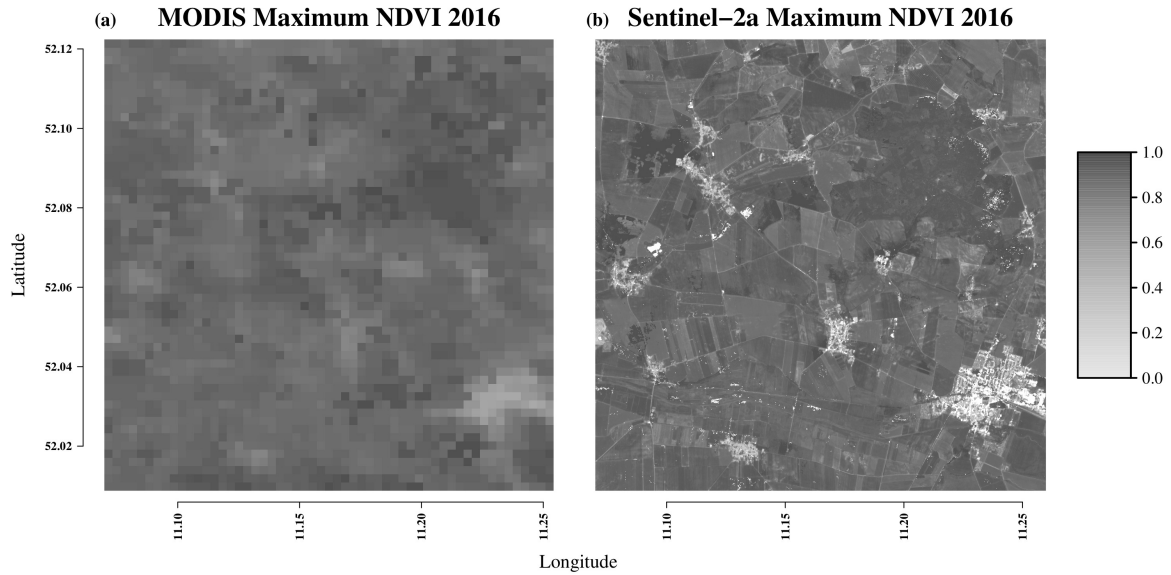


Figure 1.1: Figure (a) and (b) display MODIS Terra and Sentinel-2a maximum NDVI values, respectively, of the area around the forest site Hohes Holz and the pasture site Grosser Bruch in 2016.

for this purpose. However, highly diffuse illumination conditions and bi-directional effects may also increase NDVI values (Lange et al., 2017), demanding for more complex filter algorithms. In addition, gap filling and data smoothing, e.g. interpolation methods, signal filters and model fitting algorithms, are commonly used in advance to phenological metrics extraction (Doktor, 2008). The latter is commonly done with thresholds methods, local extrema of the phenological curve or its derivatives, and with conceptual-mathematical models (Doktor, 2008; White et al., 2009).

The validation of LSP products is a major challenge in modern phenology research. The main issues are the before mentioned lack of sharply defined phenological metrics and the large number of methods available for extraction, the sensitivity of vegetation indices and thus of phenological metrics to sensor band configurations and uncertainties of remote sensing products due to cloud cover and other non-vegetational effects (Henebry and de Beurs, 2013). Further, validation networks or campaigns for LSP products are rare (Henebry and de Beurs, 2013). However, there are ongoing initiatives, e.g. from the Land Product Validation (LPV) Subgroup of the Committee on Earth Observation Systems (CEOS) Working Group on Calibration and Validation (WGCV, Morissette et al. (2006)) and EUROSPEC, a European network established as a COST activity (Balzarolo et al., 2011). Further, eddy

covariance (EC) sites of long-term research programs like LTER (Long Term Ecological Research, LTER (2018)), TERENO (Terrestrial Environmental Observatories, Bogen et al. (2011)) and ICOS (Integrated Carbon Observation System, ICOS ERIC (2017)) are increasingly including optical measurements in order to monitor LSP (Eklundh et al., 2011; Porcar-Castell et al., 2015).

1.3 Ground-based Phenological Measurements

Optical data sampling on the ground level provides a link between remote sensing platforms and ground observations (Eklundh et al., 2011). It enhances the understanding of satellite-based measurements by providing insights into vegetational and non-vegetational factors influencing the measured signal (Eklundh et al., 2011). Further, it plays an essential role for spatial extrapolation of vegetation-related biophysical parameters and facilitates validation and process-based modelling on regional and global scale (Balzarolo et al., 2011). In case of phenology research, near-surface sensor-derived products are bridging the gap between individual-based phenology records to LSP products (Richardson et al., 2013; Lange et al., 2017).

The coordination of in-situ spectral sampling sites requires a network with unified set-up, calibration status and measurement protocols in order to allow comparability between sites (Porcar-Castell et al., 2015). Several networks and initiatives deal with these issues, e.g. Fluxnet (FLUXNET, 2017), SpecNet (Gamon et al., 2006b, 2010) and EUROSPEC (Balzarolo et al., 2011). Measurements commonly applied span across a variety of spectral, spatial and temporal resolutions, depending on the aim of the analysis (Eklundh et al., 2011). LSP products are typically acquired by the use of radiometric (Balzarolo et al., 2011; Porcar-Castell et al., 2015), e.g. multi- or hyperspectral sensors, and imaging sensors, e.g. digital cameras (Richardson et al., 2007, 2013; Eklundh et al., 2011). However, digital cameras lack a common standard for phenological research and absolute calibration is rarely done, thus signal and spectral drifts are difficult to quantify (Eklundh et al., 2011; Richardson et al., 2013). They should be treated as a complementary tool to radiometric sensors when monitoring phenology (Eklundh et al., 2011). Non-imaging radiometric sensors output a single number per measurement, integrating over the sensors field of view (FOV, Richardson et al. (2013)). These sensors can be distinguished by their FOV and their spectral sensitivity - both can range from narrow to wide (Richardson et al., 2013). A measurement taken with

wide FOV of 180° is typically referred to as hemispherical measurement, whereas narrow FOV measurements are referred to as conical or, in special cases, directional measurements (Nicodemus et al., 1977; Eklundh et al., 2011; Porcar-Castell et al., 2015). Broadband sensors are commonly used to measure photosynthetic photon flux density or total shortwave solar radiation, whereas narrowband sensors target specific regions of the electromagnetic spectrum and may be used to establish relationships between the physiology of plants and reflected radiation or to compare ground-based measurements with satellite data by using the very same band configurations (Richardson et al., 2013). Radiometric sensors with few bands are typically referred to as multispectral sensors (e.g. SKR1850 from Skye Instruments Ltd., Llandrindod Wells, Powys, U.K.), and radiometric sensors able to measure hundreds of bands simultaneously are commonly referred to as hyperspectral sensors (e.g. QE65000 from Ocean Optics, Dunedin, FL, USA or FieldSpec 4 from Analytical Spectral Devices Inc., Boulder, CO, USA) (Richardson et al., 2013). Radiometric sensors typically require a wavelength calibration, ensuring that the band positions are well characterised, and a radiometric calibration, converting the detector units into physical units (Balzarolo et al., 2011). Since the sensors degrade over time and are sensitive to temperature, the stability of the calibrations has to be checked on a regular basis (Porcar-Castell et al., 2015). Using these sensors under outdoor conditions in long-term continuous field measurement set-ups introduces error sources besides the known issues: The sensors are exposed to weather conditions and wild animals, potentially accelerating the degradation of housings, optic elements, electronics and mechanical parts or causing damage.

In-situ spectral measurements commonly involve the measurement of up- and downwelling radiation fluxes conducted with variable set-ups, including one or two sensors, single or dual FOV (SFOV, respective DFOV), different viewing geometries and varying targeted reflectance factors (Balzarolo et al., 2011; Porcar-Castell et al., 2015). Each set-up shows advantages and disadvantages and will depend mostly on the purpose of the measurement, the characteristics of the site and the research budget (Porcar-Castell et al., 2015).

1.4 Study Area

The work presented in this study deals with different scales: Whereas the first study analyses data from phenological and meteorological observation stations distributed over Germany in central Europe, the second study deals with satellite observations on a European scale. The

third and fourth study uses data from study sites situated in central Germany near the Harz Mountains (see Figure 1.2) in a temperate climate.

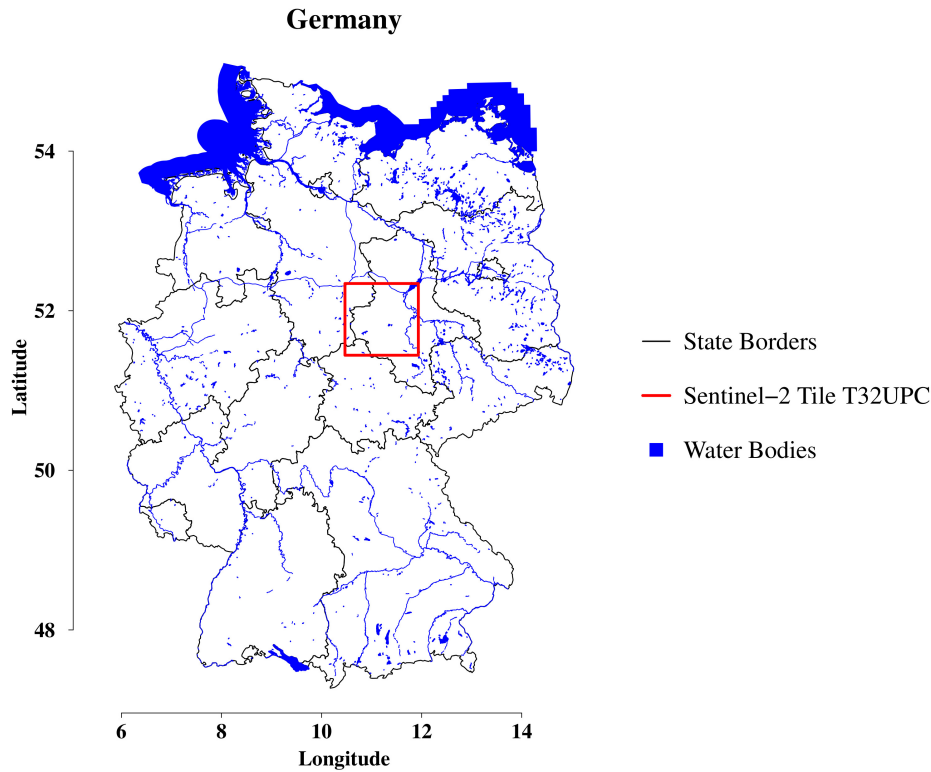


Figure 1.2: Map of Germany indicating the study site location, including state borders (Bundesamt für Kartographie und Geodäsie, 2017) in black and water bodies extracted from the CORINE dataset (Bundesamt für Kartographie und Geodäsie, 2012) in blue. The location of the Sentinel-2 tile T32UPC related to the study sites is marked in red.

© GeoBasis-DE / BKG 2017

Europe covers an area of around ten million square kilometres, the mainland spans from 36° to 71°N and 9°W to 67°E. The main part of Europe is situated in a temperate climate (see Figure 1.3), whereas southern parts are located in the subtropical and northern parts in the circumpolar zone (Köppen, 1918; Stanners and Bourdeau, 1995). The western parts are dominated by humid westerly winds, whereas the eastern parts show a continental climate (with greater diurnal and seasonal temperature variations, Köppen (1918); Stanners and Bourdeau (1995)). Germany is situated in the transition zone between both climates (see Figure 1.3), spanning from 47° to 54°N and 6° to 15°E. Coastal areas and the North-west exhibit an oceanic climate, whereas the climate is more continental in the inland zone and eastern parts (Köppen, 1918; Müller-Westermeier, 2002). Further, increasing altitude

from North to South modifies these general climate patterns, resulting in considerably different climate conditions (e.g. temperature and precipitation) throughout Germany (Doktor, 2008). Here, satellite data from AVHRR covering Europe (see Chapter 3), phenological observations extracted from the plant phenological online database (PPODB, Dierenbach et al. (2013)) covering phenological stations all over Germany observing from 1952 to 2009 (see Chapter 2), as well as temperature data measured by stations distributed all over Germany from the German Weather Service (Deutscher Wetterdienst, DWD) and daily mean temperatures obtained from climate scenarios processed by the Climate Office for Central Germany at the Helmholtz-Centre for Environmental Research - UFZ (see Chapter 2) were used.

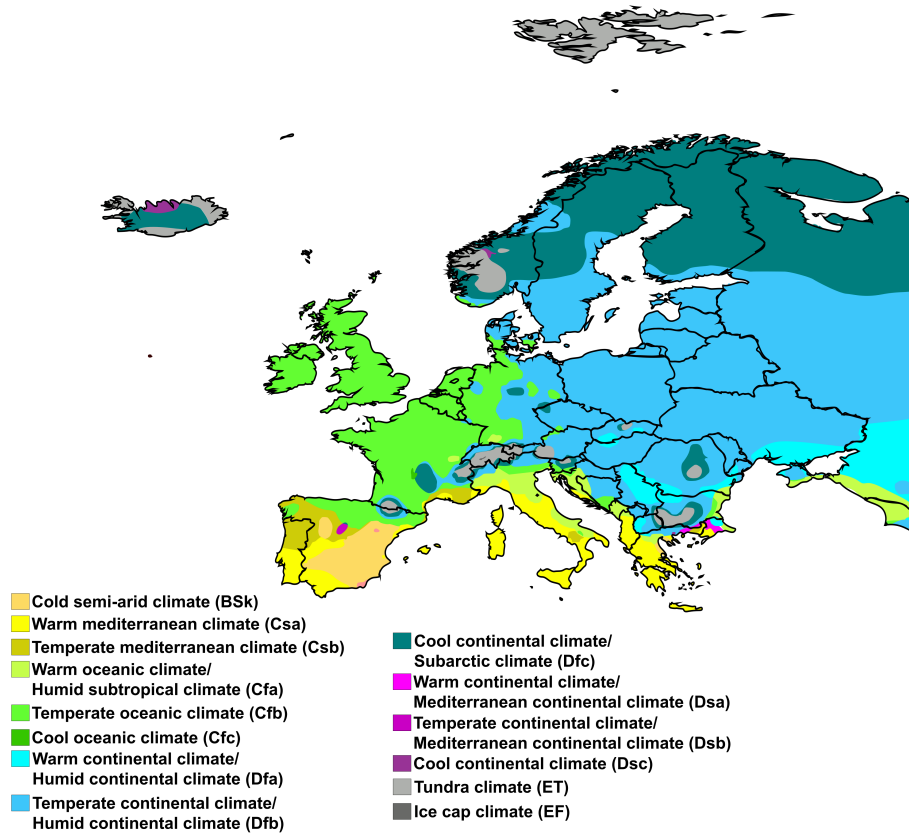


Figure 1.3: Europe map of Köppen climate classification from Peel et al. (2007). Colours indicate the different climate types: arid (type B) in brown, temperate (type C) in yellowish and greenish colours, cold (type D) in bluish and purplish colours and polar (type E) in greyish colours.

The two study sites, "Hohes Holz" and "Großer Bruch", are located in central Germany in the Mägebürger Börde, north-east of the Harz Mountains (see Figure 1.4). The two

sites are part of the Bode Catchment from the TERrestrial ENvironmental Observatories (TERENO, Wollschläger et al. (2016)). The research site "Hohes Holz" (52.08° N, 11.22° E) is a mixed deciduous forest site equipped with an eddy flux tower with a height of around 50m (Wollschläger et al., 2016). The forest itself covers an area of about 15 km² and extends about 4 km in south-north and 5 km in east-west direction. The research site "Großer Bruch" (52.03° N, 11.01° E) is a grassland site used as meadow within a nature protection site surrounding the water channel Großer Graben and is usually flooded naturally once or twice a year (Wollschläger et al., 2016). The lowland itself has wetland influences and covers an area of about 78 km² (Meynen and Schmithüsen, 1962). The site is equipped with an Eddy flux tower with a height of around 2m. Both sites are also part of ICOS (Integrated Carbon Observation System, ICOS ERIC (2017)). The towers are equipped with multi- and hyperspectral optical sensors. Digital images of surrounding vegetation are regularly taken to monitor phenology. Here, these datasets and auxiliary data taken at the sites to monitor the calibration stability of the sensors as well as for the validation of satellite phenology products from MODIS and Sentinel-2 (see Chapter 4) are used.

1.5 Research Objectives and Hypothesis

The work presented in this thesis covers the three main aspects of phenological research: forecasting phenology, thus large scale phenological modelling, the assessment of changing seasonal dynamics of vegetation on a large scale, thus remote sensing phenology products, and the validation of plant phenology products.

Forecasting phenology is a timely challenge given the necessity to predict the impact of global warming on wild-growing species and agricultural crops. Plants react to a warming climate by changing their growth patterns and reproduction habits (Kramer et al., 1996; Parmesan, 2007; Sherry et al., 2007). This may impact the environment by an enhanced vulnerability of plants to cold spells (Schaber and Badeck, 2003; Linkosalo et al., 2006) and late frost (Kramer et al., 1996; Füssel et al., 2012), consequences for species composition and distribution (Thackeray et al., 2010; Peñuelas et al., 2013) and an altered carbon exchange of ecosystems (Peñuelas et al., 2013). Although bud burst advancement was already detected in ground observational (Füssel et al., 2012; Intergovernmental Panel on Climate Change, 2014) and remote sensing data (Myneni et al., 1997; Stöckli and Vidale, 2004; Barichivich et al., 2013), it is still a matter of debate whether the response of bud burst to the predicted

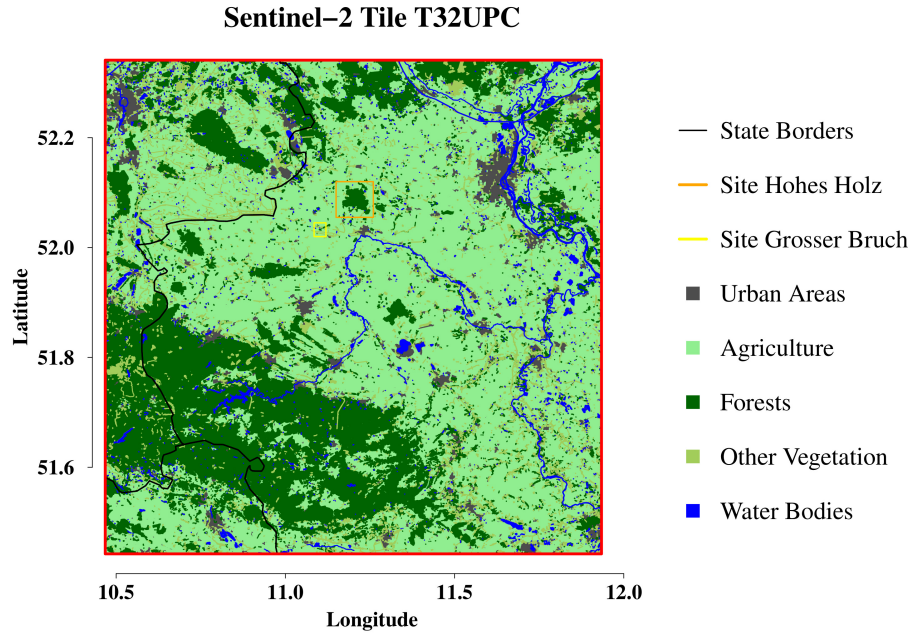


Figure 1.4: Map of the area covered by the Sentinel-2 tile T32UPC indicating the study site locations. Land cover information bases on ATKIS data (Bundesamt für Kartographie und Geodäsie, 2015). The two research sites, Hohes Holz and Grosser Bruch, are marked in orange and red, respectively. State borders are shown in black, urban areas in grey, agricultural areas in light green, forests in dark green, other vegetation in olive green and water bodies in blue.

© GeoBasis-DE / BKG 2017

climate warming will be a linear shift to earlier events. The first study of this thesis (see Chapter 2) examines whether the assumed increase of winter and spring temperatures is depicted by the phenological model PIM in correspondingly earlier bud burst, considering the complicated mechanistic interactions between chilling requirements, temperature and photo-period. Therefore, extensive measured and predicted temperature data and extensive phenological observations are used. The study tests the following hypothesis:

Hypothesis 1 *Bud burst advancement in a warming climate will be limited by species dependent chilling and photo-period requirements.*

Climate change studies utilizing large scale phenological models rely on extensive phenological and meteorological observations. Whereas the latter is now widely available, phenological observations mainly stem from human observer networks and are consequently limited

to regions with relatively high population density and thus more human observers. Satellite products may be used to obtain phenological data on a large scale even in regions where no human observers are present. One of the main challenges in remote sensing phenology products is the lack of universal definitions of phenological metrics (White et al., 2009). Numerous methods are available extracting these metrics and the various studies use different data products, varying in spatial and temporal resolution, as well as study area and period (White et al., 2009). This results in large differences between satellite derived phenological metrics and respective trend estimates (Fu et al., 2014), consequently hampering validation and inter-comparison (Henebry and de Beurs, 2013). The second study of this thesis examines various methods extracting phenological metrics. An R-package (R Core Team, 2014) for phenological data analysis called 'phenex' (Lange and Doktor, 2017) has been developed to filter the raw data, restore the seasonal NDVI pattern and to subsequently extract phenological metrics. The study tests the hypothesis that the differences between methods mainly stem from their disparate applicability for different land-use classes and mixtures thereof:

Hypothesis 2 *Methods extracting spring phenology from satellite data vary in performance depending on small scale landscape composition and thus estimate varying trends.*

Understanding the relationships between plant phenology and the remotely sensed signal at ecosystem level is crucial in order to interpret satellite-based phenology products (Fu et al., 2014; Gamon, 2015). The main challenges while establishing these relationships are different spatial resolutions, temporal mismatches between observations and non-vegetational effects influencing the remotely sensed vegetation signal (Eklundh et al., 2011). Recent satellite generations, such as the Copernicus Sentinel-2, may overcome some of these issues with their high spatial and temporal resolution and improved atmospheric correction procedures. For the third study of this thesis a validation site for Sentinel-2 products was established in central Germany. The study compares NDVI products from two multispectral sensors in DFOV mode at ground level and from a hyperspectral sensor in DFOV mode with NDVI products from satellite-borne optical sensors. Subsequently, the study examines the influence of different spatial and temporal resolutions, as well as non-vegetational effects on remotely sensed NDVI products. The study tests the following hypothesis:

Hypothesis 3 *New satellite generations, despite high temporal and spatial resolution, benefit from near-surface sensor-derived data by the provision of outlier detection and information on environmental conditions.*

During the establishment of the validation site, different problems concerning sensor calibration were encountered. Sensors in long-term continuous outdoor set-ups face harsh environmental conditions and consequently accelerated degradation (Porcar-Castell et al., 2015). Calibration shifts due to temperature and humidity as well as damage through wild animals were observed. However, calibration of these sensors and unified measurement and calibration protocols are crucial for inter-comparison with sensors mounted on other sites or satellites (Balzarolo et al., 2011; Porcar-Castell et al., 2015). The fourth study of this thesis examines the long-term stability of radiometric calibrations of automated measurements systems for multi- and hyperspectral vegetation reflectance on a continuous basis, utilising redundant measurements. The continuous calibration procedure was compared to the common campaign-based approach in order to test the following hypothesis:

Hypothesis 4 *Continuous calibration of unattended multi- and hyperspectral measurements is needed instead of campaign-based calibration procedures in order to acquire reliable measurements.*

Consequently, this thesis deals with the main challenges faced in state-of-the-art plant phenology research, covering phenological forecasting with a modelling approach, satellite-based phenology extraction, and near-surface long-term monitoring of phenology. The results presented here are thus valuable for vegetation-related climate change research, especially in the scope of agriculture, forestry and nature conservation.

2

Simulation of Forest Tree Species' Bud Burst Dates for Different Climate Scenarios: Chilling Requirements and Photo-Period May Limit Bud Burst Advancement

Authors Maximilian Lange¹, Jörg Schaber², Andreas Marx³, Greta Jäckel³, Franz-Werner Badeck⁴, Ralf Seppelt¹ and Daniel Doktor¹

Affiliations ¹ Department Computational Landscape Ecology, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany

² Institute for Experimental Internal Medicine, Otto-von-Guericke University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany

³ Department Computational Hydrosystems, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany

⁴ Council for Agricultural Research and Economics (CREA), Genomics Research Centre (GPG), Via S. Protaso 302, 29017 Fiorenzuola d'Arda, Italy

Published in: International Journal of Biometeorology 60 (11), 1711 - 1726

DOI:10.1007/s00484-016-1161-8

Publishing date: 08 April 2016

Acknowledgements The authors wish to thank the administrators of the UFZ high performance computing system, especially Ben Langenberg and Christian Krause, for their support regarding scientific computation at the UFZ in Leipzig. We would like to thank the German Weather Service (DWD) for the online provision of meteorological observations via WebWerdis and the Deutsches Klimarechenzentrum GmbH (DKRZ) for providing regional climate simulation data via the CERA database. This work was supported by the Federal Ministry for Economic Affairs and Energy Germany (grant number 50EE1218).

Author Contributions All authors contributed significantly to the work presented in this paper. Daniel Doktor and Maximilian Lange designed the present study and wrote the manuscript. Andreas Marx and Greta Jäckel prepared the climate scenario data. Maximilian Lange implemented the algorithms used in this study and analysed the data. Ralf Seppelt helped with designing the manuscript and writing. Franz-Werner Badeck and Jörg Schaber designed the model used in this study, supported its implementation and helped in designing the study. All authors edited the manuscript.

Abstract This study investigates whether the assumed increase of winter and spring temperatures is depicted by phenological models in correspondingly earlier bud burst (BB) dates. Some studies assume that rising temperatures lead to an earlier BB, but even later BB has been detected. The phenological model PIM (promoter-inhibitor-model) fitted to the extensive phenological database of the German Weather Service was driven by several climate scenarios. This model accounts for the complicated mechanistic interactions between chilling requirements, temperature and photo-period. It predicts BB with a r^2 between 0.41 and 0.62 and a RMSE of around one week, depending on species. Parameter sensitivities depict species dependent interactions between growth and chilling requirements as well as photo-period. A mean trend to earlier BB was revealed for the period 2002-2100, varying between -0.05 and -0.11 days per year, depending on species. These trends are lower than for the period 1951-2009. Within climate scenario period, trends are decreasing for beech and chestnut, stagnating for birch and increasing for oak. Results suggest that not fulfilled chilling requirements accompanied by an increasing dependency on photo-period potentially limit future BB advancement. The combination of a powerful phenological model, a large scale phenological database and several climate scenarios offers new insights into the mechanistic comprehension of spring phenology.

2.1 Introduction

A shifted onset of springtime phenological events is often detected in ground observational data (Füssel et al. 2012; Intergovernmental Panel on Climate Change 2014 and references therein). Altered phenological behaviour of plants is one of the main effects of climate change on vegetation. Environmental change causes plants to modify their growth patterns and reproduction habits (Parmesan, 2007; Sherry et al., 2007), potentially leading to consequences for species composition in ecosystems (Peñuelas et al., 2013; Thackeray et al., 2010). Phenology plays a central role in the distribution and development of species and in carbon exchange of ecosystems (Peñuelas et al., 2013). Furthermore, those species which adapt their growth behaviour quickly as a consequence of early warming are vulnerable to cold damage due to late frost, although the degree of vulnerability is different between species (Kramer et al., 1996; Füssel et al., 2012). Cold spells could also temporarily stop or even revert the development towards bud burst (BB) of late-leaving species (Linkosalo et al., 2006; Schaber and Badeck, 2003). These complex interactions possibly modify the competitive relationships between tree species (Singer and Parmesan, 2010; Thackeray et al., 2010). Besides, other global change effects, like CO₂ elevation and nitrogen deposition may impact phenology (Füssel et al., 2012; Peñuelas et al., 2013). Phenological modelling improves insight into long-term consequences of changing environmental conditions (Peñuelas et al., 2009).

The promoter-inhibitor model (*PIM*) proposed by Schaber and Badeck (2003) bases on interactions between growth promoting and inhibiting compounds in plants. These compounds are assumed to control the plants' life cycle and therefore phenology (Schaber and Badeck, 2003). *PIM* compares well with other state-of-the-art phenological models (Hänninen and Kramer, 2007; Basler, 2016) and is applicable for wider geographical regions as well as parametrised for the most important Central European tree species. BB models for tree species of former studies (e.g. Linkosalo et al. 2000; Morin et al. 2009) have been applied to datasets considering either only one species or geographically limited regions. A reason for this is the lack of insight into physiological processes leading to switches between phenological phases. Also genetic differences within species and the high intra- and inter-annual variance in timing of phenological events introduce complexity into phenological modelling. Like most present large scale phenological models for wild-growing species in temperate climates, *PIM* bases on ambient temperature and day length. These factors influence growth behaviour of plants significantly (Schnelle, 1955). Water availability is not included in most

phenological models due to the lack of data and requirement to include soil-water-plant interactions. Nevertheless, it has been found that winter precipitation could play a role for spring phenology (Fu et al., 2014).

Whether BB will respond to the predicted climate warming with linear shifts to earlier events is still a matter of debate. This is due to the complicated mechanistic interactions between chilling requirements, warming spring temperatures and photo-period (Chuine et al., 2010; Hänninen and Tanino, 2011). Some studies assume that rising temperatures lead to a corresponding lengthening of the growing season due to an earlier BB (Moser et al., 2010; Peñuelas et al., 2009), but even later BB has been detected (Morin et al. 2009; Hänninen and Tanino 2011; Pope et al. 2013).

The aim of this study was to analyse the impact of future temperature patterns on BB timing. *PIM* was implemented as R-package (R Core Team, 2014) and applied to much more extensive data than in Schaber and Badeck (2003), which differs from the former data in the length of the phenological time series, the number of phenological stations and in the availability of weather information. We implemented a fitting method reflecting new developments in optimization techniques together with a sensitivity analysis. To assess the impact of rising temperatures we used extensive climate scenario data for the years 2002 to 2100. Subsequently scenario-based phenology trends were analysed. Four species were considered: beech (*Fagus sylvatica* L.), birch (*Betula pendula* Roth.), chestnut (*Aesculus hippocastanum* L.) and oak (*Quercus robur* L.). These are the main deciduous tree species in Europe, with a high number of available phenological observations in Germany. We hypothesise that photo-period or chilling requirements could limit BB advancement even in response to future climate warming conditions.

2.2 Data and Methods

2.2.1 Temperature Data

We used two different datasets: measured data from 1951 to 2009 and climate scenarios from 2002 to 2100.

Measured temperature data (daily mean temperatures) was obtained from WebWerdis, the Web-based weather request and distribution system of Germany's national meteorologi-

cal service (Deutscher Wetterdienst, DWD). Stations providing daily mean temperatures are equally distributed over Germany. Temperature data was geostatistically interpolated to $1 * 1 \text{ km}^2$ resolution.

Daily average temperature data obtained from climate scenarios were based on two climate models, REMO-UBA (Jacob et al., 2006; Jacob and Mahrenholz, 2006) and WETTREG (Enke and Kreienkamp, 2006; Spekat et al., 2007). These data were processed in the Climate Office for Central Germany at UFZ. A bias correction using pixel-based monthly correction factors was performed after Kunstmann et al. (2004). The underlying observations provided by the German Weather Service were interpolated to the regional climate simulation grid with a resolution of $10 * 10 \text{ km}^2$. These datasets include three IPCC emission scenarios (Nakićenović et al., 2000): *A1B*, *A2* and *B1*. The scenarios are based on different story-lines describing the future development (see Table 2.1 for scenario characterisation). While *B1* assumes rapid changes towards a clean and resource efficient global society, *A2* supposes primarily regionally oriented economic development, and, consequently, higher emissions than in *B1*. The intermediate *A1B* scenario describes a future world of very rapid economic growth, including a balance across all sources of energy production.

2.2.2 Phenological Data

Phenological data were extracted from the phenological database PPODB (www.phenology.de, Dierenbach et al. (2013)). This database contains phenological observations collected by the DWD and its predecessor from 1880 to 2009. Here, we made use of data collected after 1951 to match available temperature data with a dense spatial coverage (Figure 2.1).

Phenological information is often fragmentarily available at the observing stations and contains outliers due to various sources of uncertainties, e.g. observer precision (Schnelle, 1955). Therefore, overlapping time series of adjacent stations, with a maximum distance of 5 km and a maximum elevation difference of 50 m , were combined to continuous time series (red triangles, Figure 2.1) following the methodology described in Schaber and Badeck (2002). This method estimates a yearly average day of BB from the considered time series by fitting a two-way linear model, including year and station effects. Before, outliers are removed based on a robust fit. The R-package *pheno* (Schaber, 2012) was used for the combination process. At least 5 stations are used to create one combined time series. Hereafter, this dataset is

Dataset		T (°C)			T _Σ (°C)	
		mean	sd	sd(mean _{yearly})	mean	sd
1951-2009	(reference)	8.22	7.44	0.62	496	182.4
Scenarios		Δ mean	Δ sd	Δsd(mean _{yearly})	Δ mean	Δ sd
REMO-UBA	A1B	3.44	−0.37	0.16	418.8	40.9
	A2	3.29	−0.39	0.22	400.1	62.4
	B1	2.31	−0.38	0.10	244.9	19.4
WETTREG A1B	medium	2.39	−0.53	−0.11	339.0	14.5
	wet	2.31	−0.56	−0.12	341.9	26.7
	dry	2.50	−0.32	−0.13	344.5	0.4
WETTREG A2	medium	2.30	−0.45	0.00	324.7	15.2
	wet	2.30	−0.47	−0.15	316.7	12.9
	dry	2.15	−0.31	−0.09	307.5	8.7
WETTREG B1	medium	1.86	−0.42	−0.05	252.1	2.2
	wet	1.97	−0.38	−0.04	244.5	−1.0
	dry	2.05	−0.42	−0.07	274.4	10.4

Table 2.1: Characterisation of temperature datasets. Mean and standard deviation of temperature T and standard deviation of yearly mean T as well as mean and standard deviation of temperature sums T_{Σ} are given for temperature observations (1951-2009). The differences between observations and the statistics for T and T_{Σ} of climate scenarios period 2060-2090 are shown below.

called $data_{cts}$.

In order to generate a representative dataset for the fitting procedure, one time series per natural land unit group (Meynen and Schmithüsen, 1962) was chosen (blue circles, Figure 2.1). These natural land unit groups are morphological, orographical and vegetational homogeneous regions. Furthermore, every second year was disregarded due to computational runtime purposes in the fitting procedure. These time series are henceforth referred to as $data_{fit}$. All excluded stations and years ($data_{cts}$ without $data_{fit}$) are used as validation dataset $data_{val}$. The preprocessing methodology is freely available as R-package *phenmod* (Lange, 2013).

$data_{cts}$ exhibits a mean intra-annual variance of 48 days for beech, 60 days for birch, 67 days for chestnut and 66 days for oak. The mean bud burst occurrence shows a inter-annual species-dependent variance of 28, 53, 51 and 35 days, respectively (see Figure 2.2a-d).

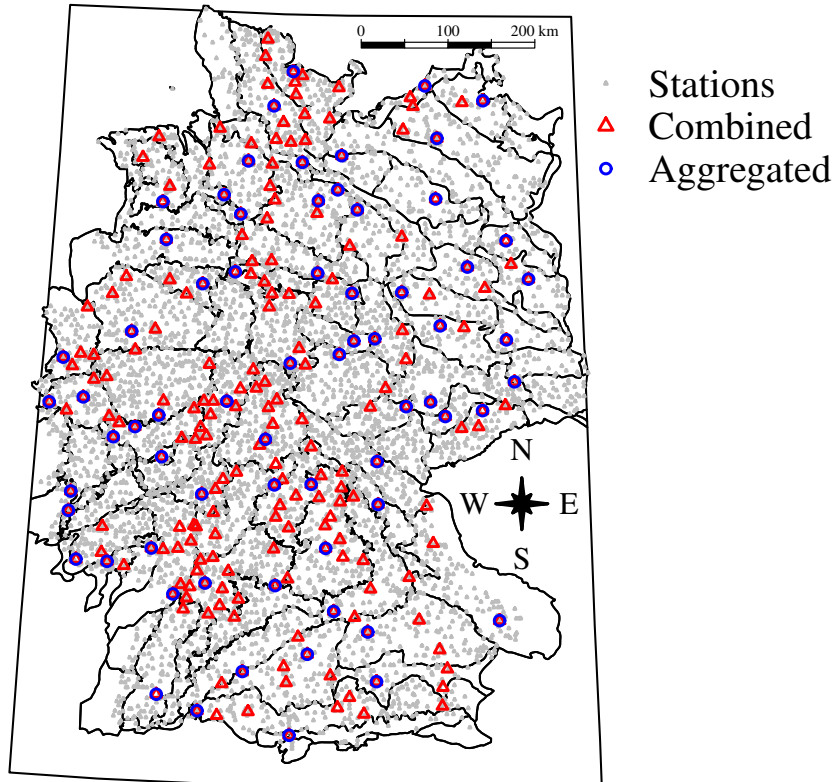


Figure 2.1: Phenological stations observing beech, birch, chestnut or oak between 1951 and 2009 (grey dots). The red triangles show the position of combined time series. The blue circles display the position of the chosen combined time series after aggregation process.

2.2.3 Promoter-Inhibitor-Model

The *promoter-inhibitor-model* (*PIM*) introduced by Schaber and Badeck (2003) is a generalized physiology-based model (Hänninen and Kramer, 2007) based on interactions between generic compounds controlling bud development. While synthesis of proteins is mainly driven by temperatures in a certain range (Johnson and Thornley, 1985; Bonhomme, 2000; Saxe et al., 2001), the phytochrome system of plants reacts to different illumination conditions and therefore photo-period (Wareing, 1956; Wareing and Saunders, 1971; Basler and Körner, 2014). These basic principles led to the formulation of *PIM*.

PIM considers two types of compounds, growth promoting compounds P and growth inhibiting compounds I , both take values between 0 and 1. The amount of these compounds

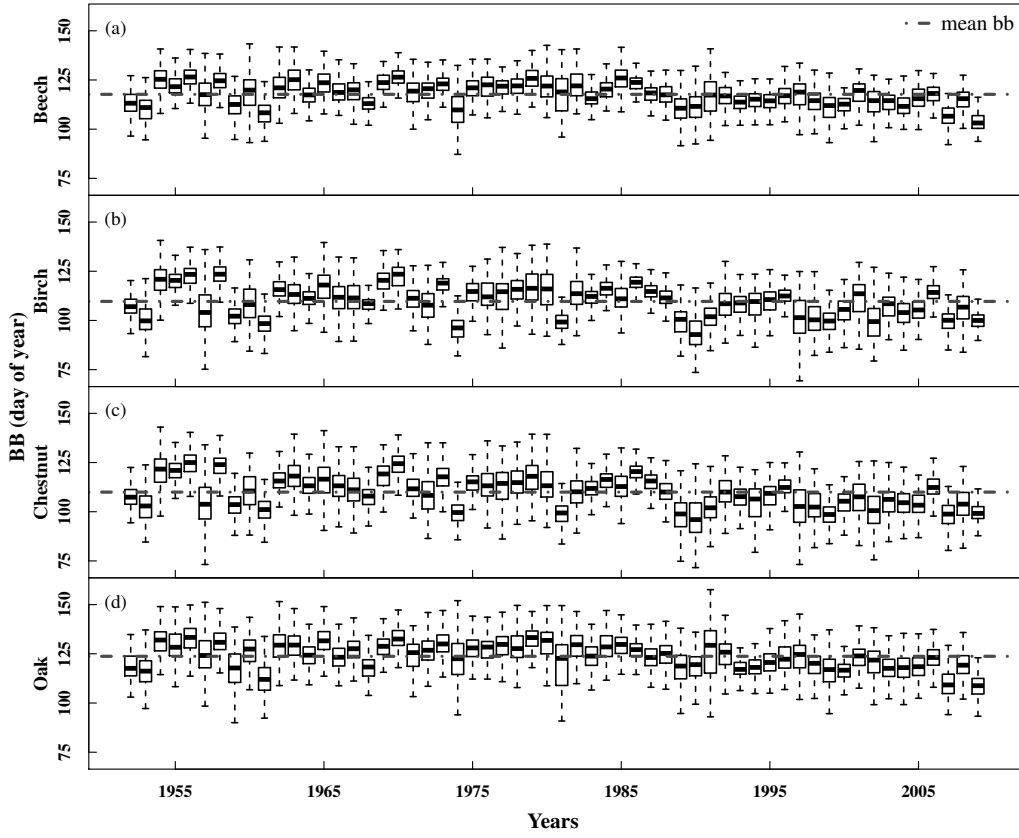


Figure 2.2: Intra-annual variance of phenological observations. The graphs are showing data of beech (a), birch (b), chestnut (c) and oak (d). The abscissa illustrates the observation years, whereas the ordinate depicts the bud burst date as Julian day.

mutually influences each other, and values are calculated in discrete steps of one day according to:

$$\begin{aligned} \Delta I &= a_1 * f_1(T) * g_1(L) \\ &\quad - a_2 * f_2(T) * g_2(L) * I \end{aligned} \quad (2.1)$$

$$\begin{aligned} \Delta P &= a_3 * f_3(T) * g_3(L) * (1 - I) \\ &\quad - a_4 * f_4(T) * g_4(L) * P \end{aligned} \quad (2.2)$$

with initial values $P(t_0) = 0$ and $I(t_0) = 1$. Function $f_i(T)$ introduces temperature dependency and function $g_i(L)$ introduces photo-period dependency (see below). t_0 is set to the previous year's day of leaf colouring if data is available and it is fixed at first of November for prediction of future bud burst to make results comparable. BB_{mod} and t_0 exhibit a nearly linear relationship. ΔP and ΔI are calculated iteratively with time steps of one day until P reaches the threshold $P = 1$, indicating bud burst (see example in Figure 2.3). The current time value t is stored as BB day. An error value BB_{NA} is produced, if P does not reach the threshold during a 200-day period.

Equation 2.3 links the daily mean temperature T of the observed region and the temperature thresholds T_{min} , T_{opt} and T_{max} regarding plant physiological processes. These thresholds are different for each i (with $i = 1, \dots, 4$), where i determines whether function $f_i(T)$ modifies the forcing ($i = 1$ or $i = 3$) or the breakdown ($i = 2$ or $i = 4$) term of the promoting ($i = 3$ or $i = 4$) or inhibiting ($i = 1$ or $i = 2$) compound.

$$f_i(T) = \begin{cases} \frac{T - T_{min,i}}{T_{opt,i} - T_{min,i}} & \text{if } T_{min,i} \leq T < T_{opt,i} \\ \frac{T_{max,i} - T}{T_{max,i} - T_{opt,i}} & \text{if } T_{opt,i} \leq T \leq T_{max,i} \\ 0 & \text{if } T < T_{min,i}, T > T_{max,i} \end{cases} \quad (2.3)$$

Equation (2.4) uses the day length L in hours. The breakdown term of the inhibiting compound and the forcing term of the promoting compound ($i = 2, 3$) are assumed to be promoted by long days, whereas the forcing term of the inhibiting compound and the breakdown term of the promoting compound ($i = 1, 4$) are assumed to be promoted by long nights.

$$g_i(L) = \begin{cases} \frac{L}{24} & \text{for } i = 2, 3 \\ \frac{24-L}{24} & \text{for } i = 1, 4 \end{cases} \quad (2.4)$$

T and L are date series with a resolution of one day. The term $1 - I$ determines that the promoting compound's development is attenuated by the value of the inhibiting compound.

Some restrictions were introduced to the Equations (2.1) and (2.2) to minimize the number of parameters. The forcing term of the inhibitor and the breakdown term of the promoter should only depend on photo-period ($f_1(T) = f_4(T) = 1$) and either a_1 or a_4 had to be zero. The temperature thresholds of $f_i(T)$ are henceforth referred to as $T_{min,I}$, $T_{opt,I}$, $T_{max,I}$ for $i = 2$ and as $T_{min,P}$, $T_{opt,P}$, $T_{max,P}$ for $i = 3$, respectively. Some terms, depending on species, should not be affected by photo-period, so $g_i(L)$ was set to 1 in these cases (see

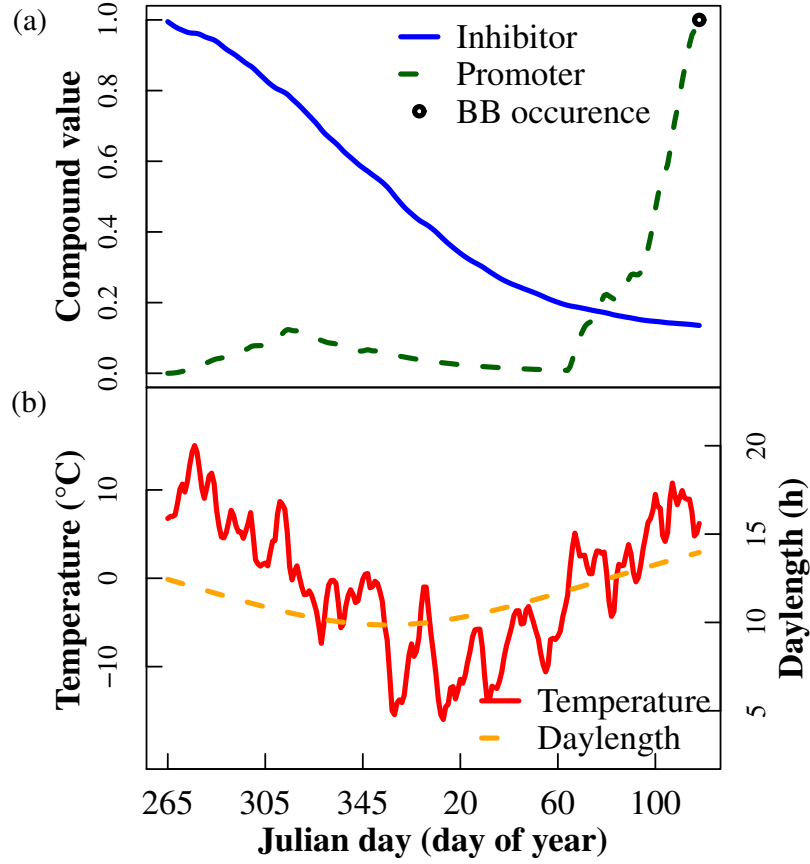


Figure 2.3: Example of development of promoter (graph a: green dashed line) and inhibitor value (graph a: blue solid line) during the model run compared with the input temperature (graph b: red solid line) and day length (graph b: orange dashed line). The abscissa shows the day of the year during the period between senescence and bud burst. BB is assumed to take place when the promoting compound's value reaches 1.

Appendix Sup.1). Chilling is included in the model structure through the degradation of the inhibiting compound at low temperatures while still accumulating the promoting compound (Schaber and Badeck, 2003). We defined the response to chilling as weighted chilling days f_{chill} in Equation 2.5, considering temperatures in the range $T_{min,P} \leq T \leq T_{max,I}$. *PIM* was implemented in the R-package *phenmod* (Lange, 2013).

$$f_{chill} = \sum_{t=t_0}^{BB_{mod}} \begin{cases} f_3(T) & \text{if } f_3(T(t)) \leq f_2(T(t)) \\ f_2(T) & \text{if } f_3(T(t)) > f_2(T(t)) \end{cases} \quad (2.5)$$

2.2.4 Fitting with DEoptim

PIM was fitted by using the R-package *DEoptim* (Ardia et al., 2011). The package provides an algorithm using *differential evolution* for optimizing a mathematical problem. Here, the objective function bases on model output (modelled BB date BB_{mod}) for a given parameter set x_k with $k = 1, \dots, j$, where j is the number of parameters (here: a_i and temperature thresholds) and observed BB date BB_{obs} . The SAR (sum of absolute residuals) of n observations was used as objective function F of *DEoptim*.

$$F = \sum_{m=1}^n |BB_{mod,m} - BB_{obs,m}| \quad (2.6)$$

DEoptim specific parameter values are shown in the supplemental material (Appendix Sup.2). Obtained parameter sets differ between each optimization run due to random initialization of *DEoptim*, therefore the procedure has to be started multiple times (here: 200 times) in order to produce robust fits. The parameter set with minimal F and maximal r^2 (Pearson, 1920) was selected as best result out of these multiple runs (see Tables 2.2 and 2.3).

Plant	a_1	a_2	a_3	a_4	$T_{opt,I}$	$T_{max,I}$	$T_{min,P}$	$T_{opt,P}$
Beech	0.000	0.018	0.185	0.050	-1.2	22.9	-9.7	28.5
Birch	0.000	0.018	0.118	0.041	-16.2	15.8	-1.0	19.5
Chestnut	0.000	0.016	0.136	0.032	-2.1	17.7	-1.1	24.2
Oak	0.002	0.008	0.457	0.000	-2.2	19.5	4.4	18.1

Table 2.2: Parameter sets retrieved from PIM fit on calibration data $data_{fit}$ for beech, birch, chestnut and oak.

Plant	r^2	MAE		RMSE	
		mean	sd	mean	sd
Beech	0.41	5.19	1.93	7.16	2.29
Birch	0.63	5.07	1.54	6.97	1.80
Chestnut	0.58	5.62	1.48	7.56	1.72
Oak	0.52	5.56	1.51	7.36	1.72

Table 2.3: Mean and standard deviation of errors and r^2 for each considered species and fitted parameter set on validation data $data_{val}$.

2.2.5 Sensitivity Analysis

A sensitivity analysis was implemented to cover two aspects: the interaction between photo-period and temperature and identifying the most influential parameters.

To assess the interaction between photo-period and temperature, we extracted 300 daily mean temperatures and day length values for each BB value in $data_{cts}$, starting at the previous year's day of leaf colouring, and used these time series as unmodified temperature and photo-period profiles. Temperature profiles were shifted with ΔT values between -12.5°C and $+12.5^\circ\text{C}$ in 0.25°C steps to create 101 different temperature profiles per unmodified time series. Photo-period profiles were shifted 101 times with ΔL values between $-3h$ to $3h$ in $0.06h$ steps. Each photo-period profile was used in combination with each temperature profile. PIM was applied to the resulting dataset containing 10201 entries per unmodified time series. Results were compared to the BB date modelled with the unmodified temperature and photo-period profile.

To identify the parameters that were most influential on the model fit, model parameters were changed *ceteris paribus* with one hundred increments in the given domain between lower and upper bound (see Appendix Sup.2). Subsequently, the SAR of the recalculated BB dates were compared in order to access information about the new parameter values' quality. A sensitivity value S_k for each parameter value x_k was introduced (see Equation 2.7).

$$S_k = \frac{\Delta SAR / SAR}{\Delta x_k / x_k} \quad (2.7)$$

S_k was calculated for valid changes Δx_k of the parameter value satisfying the equation $l_k \leq$

$x_k + \Delta x_k \leq u_k$ (with l_k as lower and u_k as upper bounds of the parameter k). Parameters are assumed to be sensitive if $S_k \geq 1$.

2.3 Results

2.3.1 Combined Phenological Time Series

Combining phenological data produces reliable merged continuous time series (see Figure 2.4), which are continuous time series which do not include extreme (local) maxima and minima. $data_{fit}$ has the following size after the above mentioned aggregation steps: 1596 data points for beech, 2002 data points for birch, 1967 data points for chestnut and 1845 data points for oak (see Figure 2.1 for respective stations).

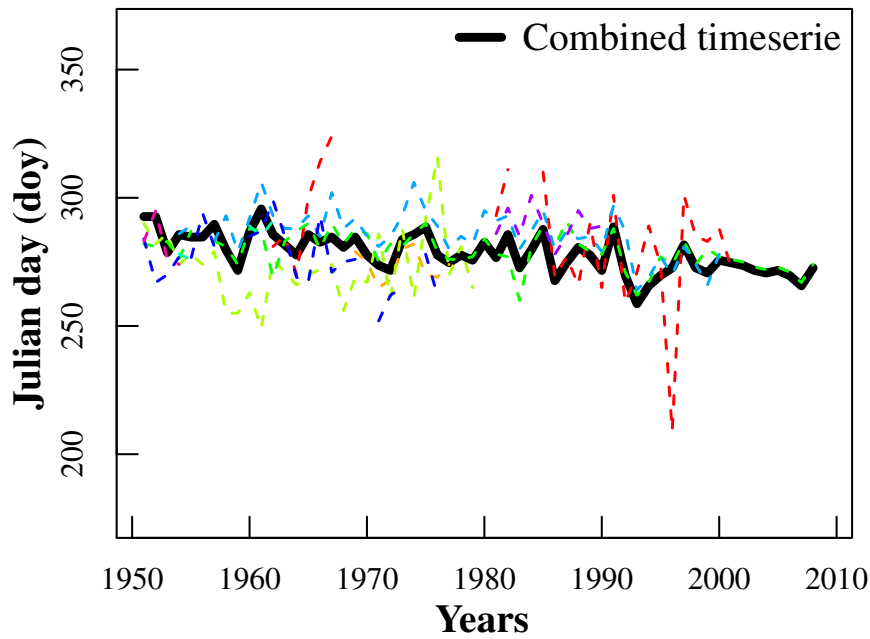


Figure 2.4: Single phenological time series (coloured and dashed lines) of leaf colouring dates and their combination (black solid line). The abscissa shows the year and the ordinate the day. Single phenological time series contain gaps and outliers, whereas the combination is a continuous time series.

2.3.2 Sensitivity Analysis and Model Fits

The effect of photo period and temperature on BB differs between species (see Table 2.4). Whereas photo-period has a strong influence on BB of beech and a moderate influence on BB of oak, it has minor effects on BB of chestnut and negligible influence on BB of birch. Temperature has strong influence on BB for birch, chestnut and oak, but only a moderate influence on BB of beech. The effect of photo-period increases with rising temperatures while the effect of temperature decreases (see Table 2.4). Weighted chilling days f_{chill} showed a strong correlation with BB (see Table 2.4).

Plant	$-12.5^{\circ} C \leq \Delta T \leq 12.5^{\circ} C$		$0^{\circ} C < \Delta T \leq 12.5^{\circ} C$		$cor(f_{chill}, BB)$
	$-3 h \leq \Delta L \leq 3 h$		$-3 h \leq \Delta L \leq 3 h$		
	$cor(L, BB)$	$cor(T, BB)$	$cor(L, BB)$	$cor(T, BB)$	
Beech	-0.73	-0.52	-0.92	0.11	-0.91
Birch	-0.11	-0.83	-0.26	0.15	-0.71
Chestnut	-0.32	-0.76	-0.64	0.18	-0.86
Oak	-0.50	-0.76	-0.78	-0.15	-0.56

Table 2.4: Correlations of photo-period L and temperature T for ΔT between $-12.5^{\circ}C$ and $12.5^{\circ}C$ on the left side and for ΔT between $0^{\circ}C$ and $12.5^{\circ}C$ in the middle. The correlations of weighted chilling days f_{chill} with bud burst are shown on the right side. All correlations are shown for beech, birch, chestnut and oak.

Nearly all model parameters are important and sensitive. The parameters a_i (with $i = 1, \dots, 4$) were analysed between 0 and 1 and most of them are highly sensitive with $S_k > 10$ on small increments. Parameter a_3 shows a high sensitivity for all species, whereas the parameters a_2 and a_4 are moderately sensitive with $S_k \geq 1$. Parameter a_1 shows only a low sensitivity with $S_k < 1$ (see Figure 2.5a-1).

Optimal promoter and inhibitor temperature values were derived between thresholds of $-24^{\circ}C$ to $+32^{\circ}C$. $T_{opt, I}$ shows low sensitivity, whereas $T_{min, I}$ and $T_{max, I}$ exhibit high sensitivity. $T_{min, P}$ shows moderate sensitivity for species beech and oak and low sensitivity for birch and chestnut. The parameters sensitive temperature ranges vary with respect to minimum and maximum thresholds. $T_{opt, P}$ exhibits high sensitivity for all species, whereas $T_{min, P}$ and $T_{max, P}$ show low sensitivity for species birch, beech and chestnut and high sensitivity for oak. $T_{min, I}$ and $T_{max, P}$ show no change in SAR beyond certain temperature thresholds and

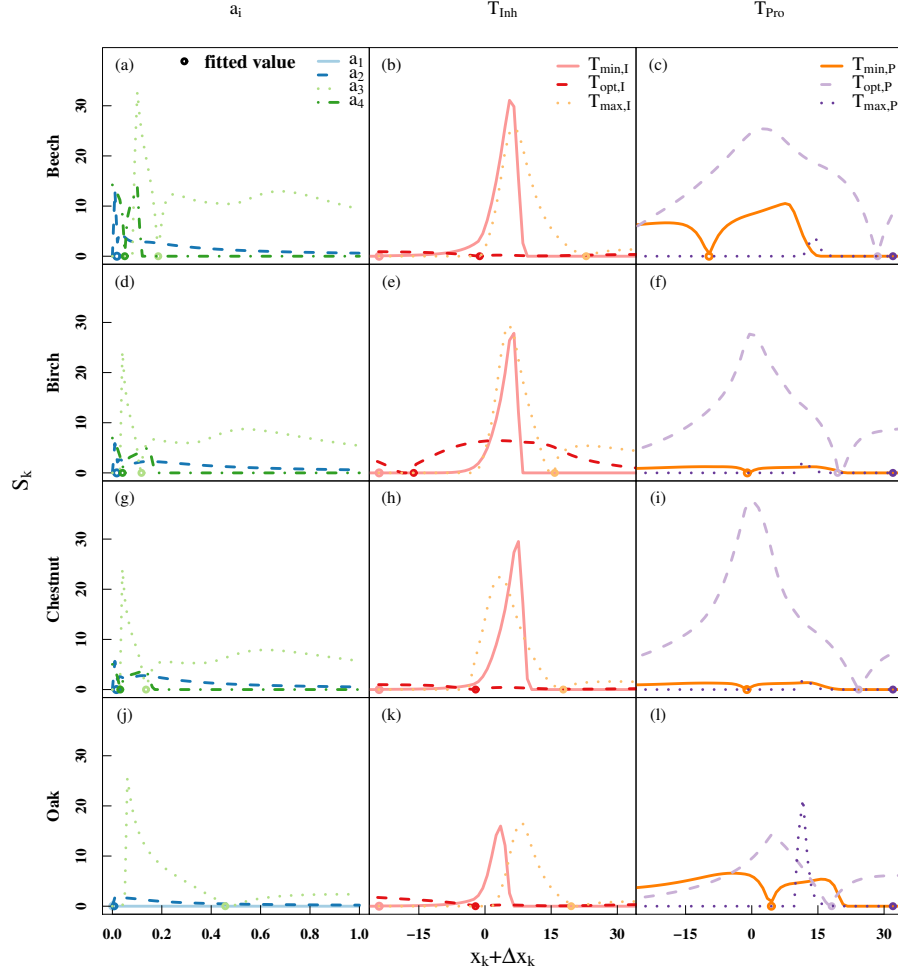


Figure 2.5: Influence of changing parameter value x_k by Δx_k on the sum of absolute residuals as sensitivity S_k of model parameters. Parameters a_i are shown on the left side (a,d,g,j) in blue and green, the inhibiting compound's temperature thresholds are shown in the middle (b,e,h,k) in reddish colours and the promoting compound's temperature thresholds are shown on the right side (c,f,i,l) in purple and orange. The fitted parameter values are shown as dots. The x-axis shows the parameter value $x_k + \Delta x_k$. The y-axis shows the sensitivity value S_k (see Equation 2.7).

were therefore fixed at $T_{min,I} = -24$ and $T_{max,P} = 32$ for further calculations.

$T_{opt,I}$ was fitted to values between -1.2°C to -2.2°C , except for birch with $T_{opt,I} = -16.2^\circ\text{C}$ (see Table 2.2). $T_{max,I}$ shows values between -15.8°C (for birch) and 22.9°C (for beech). $T_{min,P}$ exhibits fitted values between -9.7°C (beech) and 4.4°C (oak). The lowest value for

$T_{opt,P}$ is 18.1°C (oak) and the highest value 28.5°C (beech). Thus, the model is parametrised such that the inhibitor is optimally broken down at temperatures below zero. This supports the concept of chilling requirement, i.e. the need of a period of low temperatures to promote bud burst later at higher temperatures.

We found strong correlations between the parameters a_i (see Figure 2.6a-d). Parameters related to the breakdown (a_2 and a_4) term of the model are stronger correlated among themselves than to parameters of the forcing term. $T_{opt,I}$ and $T_{opt,P}$ are highly positive correlated for beech, and highly negative correlated for other species, potentially due to a more photo-period driven life-cycle of beech and therefore a higher adaptability to different spring temperatures. The parameters a_i are strongly correlated with $T_{opt,I}$ (except for birch), with $T_{opt,P}$ (beech and oak) and with $T_{min,P}$ (chestnut and oak). $T_{min,P}$ shows a strong correlation with $T_{opt,I}$ and $T_{opt,P}$ (chestnut and oak).

2.3.3 Model Performance

The fitted model was applied to the validation dataset $data_{val}$ in order to analyse model performance. The mean absolute errors (MAE) are between 5.07 and 5.62 days, the root mean squared errors (RMSE) between 6.97 and 7.56 days.

The predicted BB values BB_{mod} for birch show the lowest errors ($MAE = 5.07$ and $RMSE = 6.97$) of all considered species and the highest r^2 of 0.62. BB_{mod} for beech are estimated with the second best MAE (5.19 days) and RMSE (7.16 days), but with lowest r^2 (see Table 2.3 and Figure 2.7a,b). BB_{mod} for chestnut have the highest MAE (5.62 days) and RMSE (7.56 days). BB_{mod} for oak are estimated with a moderate r^2 , but with the second highest MAE and RMSE of all species (see Table 2.3 and Figure 2.8a,b).

2.3.4 Climate Scenarios

Mean trend estimations over all scenarios and sites show low but significant trends of BB advancement for all species, although five out of 48 scenarios show positive or insignificant trends (see Table 2.5 and Figure 2.9a-p).

PIM predicts only a weak trend towards later BB when using climate scenarios WETTREG *B1 wet* (beech, birch and chestnut) and *A2 dry* (beech). Statistics over all scenarios show

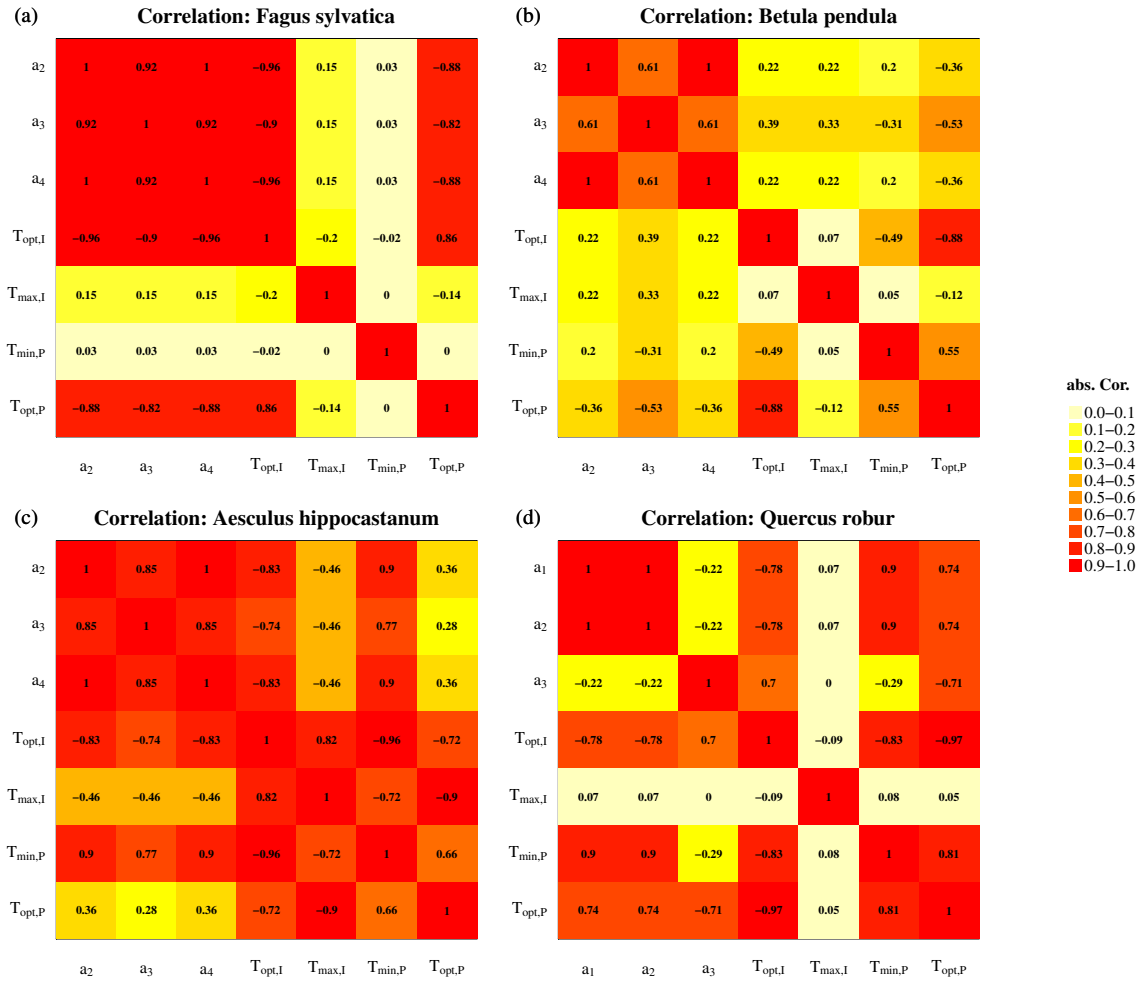


Figure 2.6: PIM parameter correlation for beech (a), birch (b), chestnut (c) and oak (d). Colours from yellow/pale (weak) over orange (medium) to red/dark (strong) show increasing absolute correlation.

a mean trend around -0.05 days per year for beech, around -0.07 d/y for birch, around -0.10 d/y for chestnut and around -0.11 d/y for oak (see Table 2.5). The strongest trend towards earlier BB is detected for REMO-UBA A2, the weakest trend for WETTREG B1 wet (see Figure 2.9a-p). Trend estimates differ more across climate scenarios than between species. REMO-UBA exhibits similar patterns across species, whereas WETTREG depicts increasing across-species differences along model assumptions *medium*, *wet* and *dry*. The number of significant trend estimates is decreasing along this gradient.

Trend statistics over all scenarios show a distinct pattern for each species when divided into

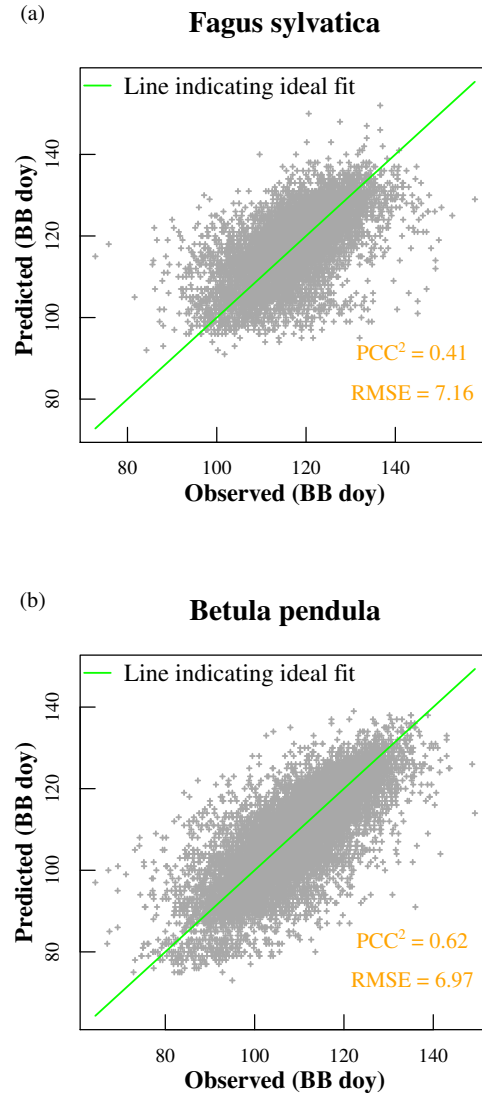


Figure 2.7: PIM predictions on *data_{cts}* compared to observed BB for beech (a) and birch (b).

two periods (2002-2050 and 2051-2100, see Table 2.5). Whereas trends towards earlier BB are decreasing for beech and chestnut, birch exhibits the same trend for both periods and oak shows a slightly stronger trend in the second period.

Furthermore, BB dates and temperature sums show a clear non-linear relationship for all climate scenarios and all species (see Figure 2.10). Temperature sums (T_{Σ}) are defined (see Equation 2.8) by the sum of daily mean temperature above 0°C (following Heide 1993).

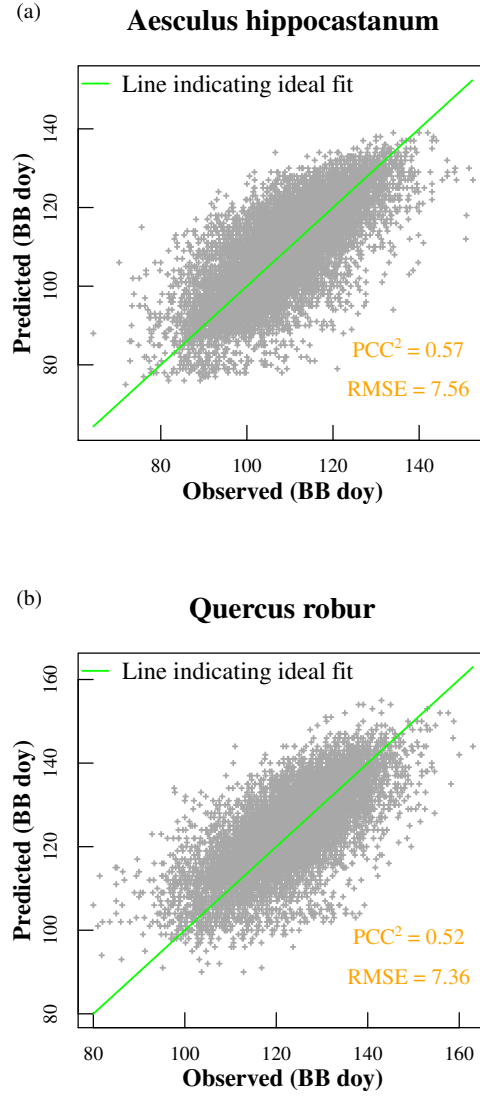


Figure 2.8: PIM predictions on $data_{cts}$ compared to observed BB for chestnut (a) and oak (b).

$$T_{\Sigma} = \sum_{t=t_0}^{t_{end}} \max(T(t), 0^{\circ}C) \quad (2.8)$$

t_0 is fixed at previous years first of November and t_{end} was set to day 120.

Slowdown of BB advancement is strongest for Beech and just moderate concerning the other species. Above a certain temperature sum threshold (around $1600^{\circ}C$, see Figure 2.10), BB

		Beech	Birch	Chestnut	Oak
1951-2009	mean	−0.13	−0.17	−0.20	−0.17
	sd	0.14	0.09	0.12	0.12
	Sign. stations	58%	65%	72%	73%
2002-2100	mean	−0.05	−0.07	−0.10	−0.11
	sd	0.02	0.02	0.02	0.02
	Sign. stations	53%	50%	73%	83%
2002-2050	mean	−0.06	−0.06	−0.10	−0.10
	sd	0.05	0.06	0.06	0.06
2051-2100	mean	−0.02	−0.06	−0.09	−0.12
	sd	0.05	0.06	0.06	0.05
	Sign. of diff.	< 0.001	0.1	< 0.001	< 0.001

Table 2.5: Trends (days per year) summarized over observation data (1951 to 2009) and all climate scenarios for the whole period (2002 to 2100) and two sub-periods (2002 to 2050 and 2051 to 2100). Mean and standard deviation of trends as well as percentage of significant stations and significance of difference of the two sub periods are shown.

occurrence of all species levels off. A BB advancement slowdown occurs at earlier temperature sums (around 500°C), with the magnitude of slowdown varying between species.

2.4 Discussion

Fitted parameter values suggest a different impact of climate warming on the four considered species. Beech exhibits a strong negative correlation between $T_{opt,I}$ and $T_{opt,P}$, whereas birch, chestnut and oak show a strong positive correlation. The fitted parameter values $T_{min,P}$ and $T_{max,I}$ reflect the adaptability to changing spring temperatures. The interval $(T_{min,P}, T_{max,I})$ is largest for beech, smallest for oak and of moderate size for the other species, suggesting a weaker impact of spring temperatures on BB of beech, a moderate impact for birch and chestnut and a stronger impact on BB of oak. Beech also exhibits the strongest impact of photo-period on BB. The BB sensitivity of beech to photo-period is consistent with earlier modelling results (Schaber and Badeck, 2003) and observations of Basler and Körner (2012). The correlation of photo-period and BB is moderate for oak and minor for chestnut and birch, although Myking and Heide (1995) and Caffarra et al. (2011) reported photo-period sensitivity in *Betula* when chilling was incomplete. The distinct and

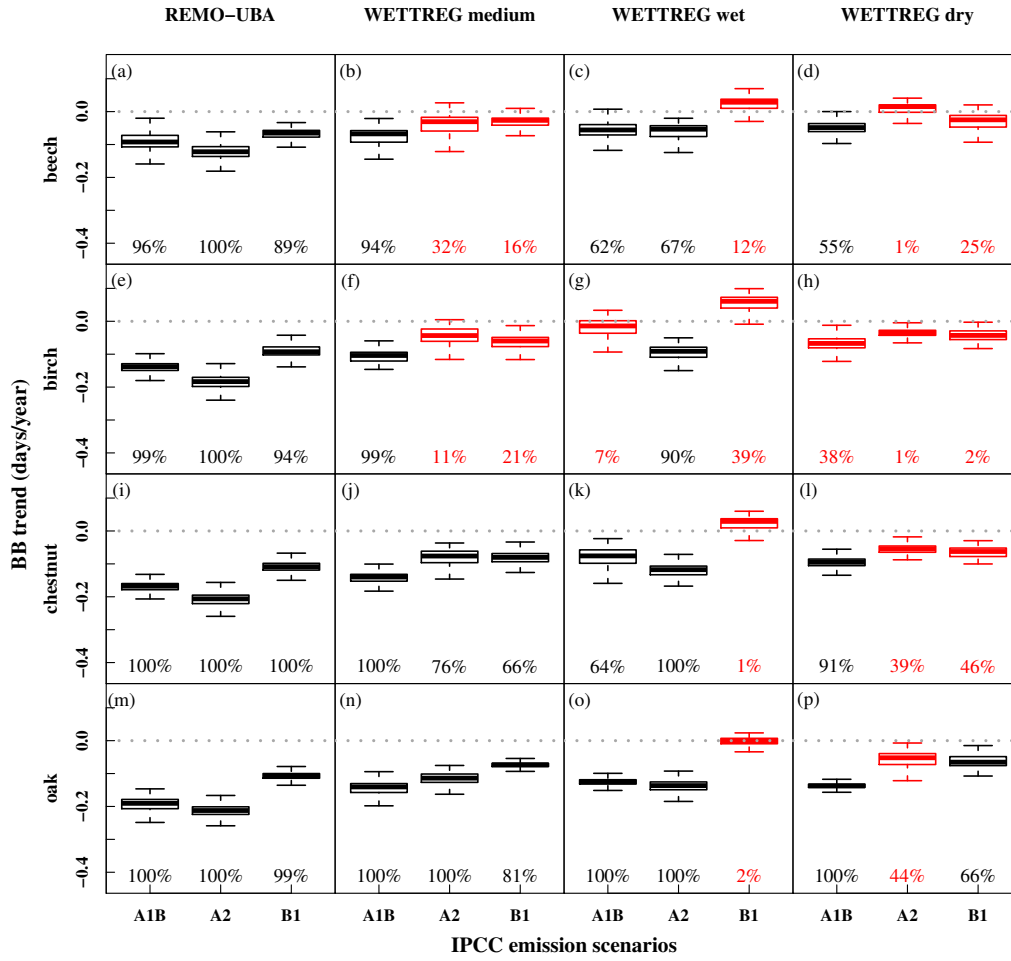


Figure 2.9: Trends of beech (a-d), birch (e-h), chestnut (i-l) and oak (m-p) for period 2002-2100 are summarized over all twelve climate scenarios. All non-red box-plots are significant. The percentages show the number of stations with significant trends.

varying interactions of photo-period and temperature consequently denote a species-specific climate change impact. Whereas beech shows a strong attenuation of BB advancement under future climate warming conditions, birch and chestnut exhibit a moderate and oak a minor slowdown. Respective correlations of BB with f_{chill} and the impact of spring temperatures on BB support these results, which are reflected by trend estimates: decreasing for beech and chestnut, stagnating for birch and increasing for oak. The increasing impact of photo-period on BB, specifically within high temperature regimes, strongly points to unfulfilled chilling

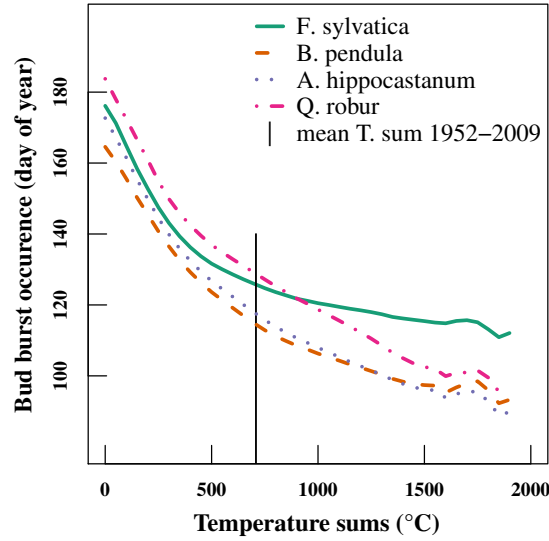


Figure 2.10: Relationship of BB development and temperature sums. The modelled BB is averaged for all climate scenarios. BB of beech is displayed as green solid line, birch as brown dashed line, chestnut as blue dotted line and oak as pink dotted and dashed line. The vertical line displays the mean temperature sum of temperature observations between 1952 and 2009.

requirements. These results are in line with those reported by Laube et al. (2014). Temperature responses, reflected by the temperature thresholds T_{min} , T_{max} and T_{opt} , differ between species. Whereas temperature effects for beech level out, birch and chestnut show a low sensitivity. Oak exhibits the highest $T_{min, P}$ and therefore highest heat requirements, leading to a high impact of increasing spring temperatures on BB. The diverging trend behaviours reveal that trend attenuation occurs at species specific levels of temperature increase and may not be encountered in the period 2002-2100. The extreme temperature increase simulated in the modelling parameter analysis (see Table 2.4) is not completely covered by the climate scenario temperature data, but may occur under future climate conditions beyond the analysed period.

PIM reacts highly sensitive to parameter variations due to their high correlation. As suggested by Schaber and Badeck (2003), we fixed the parameters $T_{min, I}$ and $T_{max, P}$ due to their low sensitivity in physically meaningful ranges (below -10°C respective above 20°C) and their high correlation with other parameters. The parameters a_i regulate the interaction of growth promoting and inhibiting compounds' degradation and synthesis. Hence, they are most sensitive. Temperature thresholds are not as sensitive as the parameters a_i , possibly due

to the adaptability of plants to temperature ranges. Extrema are different for each considered species, indicating potential differences in the optima of their temperature response.

Fu et al. (2015a,b) found the temperature sensitivity of leaf unfolding was significantly reduced over 1980-2013, as heat requirements were increased. The authors identified that reduced winter chilling increased heat requirements and also an uncertain influence of the photo-period (Keenan, 2015). This is in line with our results, which suggest that unfulfilled chilling requirements lead to an increasing photo-period impact and therefore a reduction of temperature sensitivity, reflected in an increased heat requirement.

Model errors are of low magnitude considering the high natural intra- and inter-annual BB variance of several weeks (Doktor, 2008; Schaber and Badeck, 2002). They compare well to other studies modelling BB using different datasets and models (Linkosalo et al., 2000; Morin et al., 2009; Vitasse et al., 2011; Fu et al., 2012; Olsson and Jönsson, 2014; Basler, 2016). Mean trend estimations of predicted BB are in line with other studies (Vitasse et al., 2011; Migliavacca et al., 2012). Errors and predictions mainly differ in consequence of different topologies and sizes of observation areas with a resulting difference in natural phenological variability.

The mean trend of all climate scenarios is -0.05 up to -0.11 days per year, depending on species. Trend estimates for *WETTREG* under the scenarios A2 and B1, characterised by lesser spring temperature increase and a decreased temperature variability, exhibit strong across-species differences and a high number of insignificant trends. Trends are predominantly significant for all other scenarios, although they are of low magnitude considering the previously mentioned high natural intra- and inter-annual BB variance. Despite this, trend estimates are reliable. The methodology used removes potential error sources by eliminating outliers and gaps in time series and therefore decreasing variance, enabling a robust depiction of BB in a changing climate. These time series still contain a high natural variability of phenological events.

The Intergovernmental Panel on Climate Change (2014) expects an earlier onset of spring phenological events in humid temperate and boreal forests, with some species showing a reduced response to global warming due to photo-period and chilling requirements. Our results suggest a slowdown of BB advancement for most considered species in the future. These results are in line with Morin et al. (2009). All modelled trends decreased between the observation period 1951-2009 and the scenario period 2002-2100. There is a nearly linear relationship between rising temperature sums and earlier BB, but BB advancement will slow-

down beyond 500°C (see Figure 2.10). Oak exhibits a minor slowdown of BB advancement and is the only species with a distinctive trend for all climate scenarios and an increasing trend compared to the results of Schaber and Badeck (2005).

2.5 Conclusions

PIM depicts a species dependent impact of photo-period on BB, especially for beech. Results assume a stagnation of BB advancement due to limited light level and unfulfilled chilling requirements, as mentioned by other studies (Hänninen and Tanino, 2011). This effect may not occur in this century for all species due to their different temperature responses, leading to consequences in species composition of ecosystems. Hänninen and Tanino (2011) and Chuine et al. (2010) assume a species dependent non-linear response to global change due to unfulfilled chilling requirements. Consequently, spring warming would lead to a growing season lengthening as long as temperatures will not exceed a species specific temperature sum limit and photo-period requirements are fulfilled. This hypothesis is supported here by the detection of lower trends in future time series compared to past time series.

The results of our work provide new insights into the mechanistic comprehension of spring phenology. According to this study, considered plant species react differently to changing future environmental conditions. This corresponds with their varying dependency on photo-period and temperature and may alter the competitive balance between species. Further research on this topic should test non-linear dependencies between temperature sums and phenological state. Future studies have access to the excellent phenological database PPODB and, due to this work, methods for preprocessing phenological and temperature data and for phenological modelling.

Supporting Information

Sup.1 Selected Promoter-Inhibitor-Models

The best model for each selected species was chosen, therefore four models were used. For species birch (*B. pendula*), model number two with $a_1 = 0$ and no photo-period dependence for the forcing term of promoter and the breakdown term of inhibitor was used:

$$\begin{aligned}
\Delta I &= -a_2 * f_2(T) * I \\
\Delta P &= a_3 * f_3(T) * (1 - I) \\
&\quad - a_4 * \frac{24 - L}{24} * P
\end{aligned} \tag{2.9}$$

For species chestnut, model number eight with $a_1 = 0$ and no photo-period dependence for the forcing term of promoter was used:

$$\begin{aligned}
\Delta I &= -a_2 * f_2(T) * \frac{L}{24} * I \\
\Delta P &= a_3 * f_3(T) * (1 - I) \\
&\quad - a_4 * \frac{24 - L}{24} * P
\end{aligned} \tag{2.10}$$

For species beech, model number eleven with $a_1 = 0$ was used:

$$\begin{aligned}
\Delta I &= -a_2 * f_2(T) * \frac{L}{24} * I \\
\Delta P &= a_3 * f_3(T) * (1 - I) * \frac{L}{24} \\
&\quad - a_4 * \frac{24 - L}{24} * P
\end{aligned} \tag{2.11}$$

For species oak, model number twelve with $a_4 = 0$ was used:

$$\begin{aligned}
\Delta I &= a_1 * \frac{24 - L}{24} - a_2 * f_2(T) * \frac{L}{24} * I \\
\Delta P &= a_3 * f_3(T) * (1 - I) * \frac{L}{24}
\end{aligned} \tag{2.12}$$

Sup.2 DEoptim Parameter Settings

The optimization with DEoptim was done with following parameters:

2 Modelling Phenology

- population size $NP = 400$
- iteration number $It = 100$
- crossover rate $CR = 0.9$

These parameters determine the fits' convergence behaviour. The parameter bounds were set to following values:

- -24° to 32° Celsius for temperature thresholds
- 0 to 1 for scaling parameters a_i

3

Mixed pixel effects and land-cover type translate method dependent into varying spring phenology trend estimates using coarse resolution satellites

Authors Daniel Doktor^{1,*}, Dirk Koslowsky², Maximilian Lange¹, Ralf Seppelt¹ and Franz-Werner Badeck³

Affiliations ¹ Department Computational Landscape Ecology, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany

² Free University Berlin, Department of Earth Sciences, Institute of Meteorology, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany

³ Council for Agricultural Research and Economics (CREA), Genomics Research Centre (GPG), Via S. Protaso 302, 29017 Fiorenzuola d'Arda, Italy

* Correspondence: daniel.doktor@ufz.de; Tel.: +49-(0)341-235-1943

Submitted to: Remote Sensing of Environment

Submission date: 16 May 2019

Acknowledgements This project was supported by the Federal Ministry for Economic Affairs and Energy Germany (grant number 50EE1218). The authors would like to thank especially Ben Langenberg and Christian Krause for providing and maintaining the UFZ high performance computing system. This study greatly benefited from ground based phenological measurements provided by the Deutscher Wetterdienst (DWD) and its numerous (honorary) observers.

Author Contributions Daniel Doktor designed the study, performed the radiative transfer simulations as well as phenological metrics extraction and wrote the paper. Maximilian Lange and Daniel Doktor designed the methods and data structures used for phenological metrics extraction and implemented them into the R-package 'phenex'. Maximilian Lange set-up the framework for the phenological metrics extraction on a high performance computing cluster, as well as download, pre- and post-processing procedures for large amounts of satellite data which were used to compare NOAA AVHRR results to NASA-MODIS-based phenological metrics. Franz-W. Badeck, Maximilian Lange and Ralf Seppelt edited the manuscript. Dirk Koslowsky provided the inter-calibrated NOAA AVHRR data.

Abstract Various studies assessed the vegetation response to modified temperature patterns by extracting phenological metrics from satellite observations. While assessments based on ground observations show a consistent trend towards earlier season onset across species, corresponding studies relying on remotely sensed imagery reveal a more diverse picture. When averaging at continental scales mostly negative trends (towards earlier onset) have been identified but trend magnitudes varied considerably. At land-cover level both negative and positive trends have been found also at comparably small spatial scales. We hypothesise that these apparent trend differences between land-cover types stem from mixed pixel effects which translate method dependent on extracted phenological metrics. These are more pronounced when employing coarse spatial resolution satellite imagery (> 1 km) where mixed green-up signals from different land-cover types are common.

Therefore, this work analysed first – based on simulated satellite observations with implemented known trend to earlier green-up – how a mixture of different green-up signals potentially translates into varying extracted metrics and respective trends. Here, the most commonly used methods to extract green-up dates were implemented. Secondly, these methods were tested across land-cover types based on real data. NOAA AVHRR daily NDVI observations of 1 km spatial resolution (1989-2007) covering most of Europe were used. Extracted

green-up days and trends were evaluated against an extensive phenological ground network and also a phenological model for the area of Germany.

The implemented framework allows to assess the robustness of derived green-up trends stemming from heterogeneous landscapes on a theoretical basis. Results from simulated satellite data suggest that varying trend estimates for many landscapes originate from mixed green-up signals, i.e. two or more land-cover types within one pixel and are even found within homogeneous land-cover. Derived trends strongly diverge from the implemented trend especially in case of mixed signals. Regarding real satellite data, phenological metrics vary across methods and land-cover classes. Respective mean trend estimates differ between 0.5 – 1.5 days/year towards an earlier greening up but also include more negative or even positive trends. Mean trend estimates are in line with observed trends on the ground and modelled trends for the area of Germany.

Trend differences within and across land-cover classes can be explained based on the presented theoretical framework: diverging performance of some methods – especially in heterogeneous landscapes – where the satellite often captures a mixture of green-up signals due to the satellite's coarse spatial resolution, crop rotation and land-cover change.

3.1 Introduction

Phenology is the study of recurring life cycle events mostly related to biological processes which trigger for example the leaf-out or leaf-colouring of trees. The timing of these events can differ substantially between regions but also between years for a single region and provides a measure of an ecosystem's response to variable and changing temperature and precipitation patterns. Changes in vegetation phenology affect carbon, water and energy exchange between vegetation and atmosphere. The phenological timing, for example, defines the carbon uptake period of vegetation (Baldocchi, 2003). It also substantially modifies energy exchange patterns in the lower atmospheric boundary layer (Schwartz and Crawford, 2001) which is of importance for many climatological models.

With respect to plant species distribution and composition most of the recent research concludes that those parts of a plant population which respond quickly to a comparably early warming up risk damage through cold spells or late frost, albeit to different degrees dependent on the species (Hänninen, 1991; Repo et al., 1996). Changing temperature regimes in spring time can alter competitive relationships between tree species (Kramer, 1994, 1995).

3 Satellite Phenology Products

Considering fauna-flora interactions and dependencies there has been evidence of i) increased stress for invertebrates who rely on synchrony with their hosts (Visser and Holleman, 2001; Parmesan, 2007), ii) temporal changes in predator-prey synchrony (Visser and Both, 2005) and population decline of migratory birds (Saino et al., 2011).

Ground observational networks consistently report a changing phenological timing from the 1980s onwards, which, for example, results in longer vegetation periods (e.g. Keeling et al. (1996); Menzel and Fabian (1999); Fitter and Fitter (2002); Schwartz et al. (2006)). It is commonly agreed that these changes reflect the global warming signal (Parmesan and Yohe (2003); Root et al. (2003); Walther (2004); Root et al. (2005); Schaber and Badeck (2005)). Natural variability and observer error affect the robustness of ground recordings (e.g. bud burst day, d_{BB}). But usually these recordings are compared to satellite derived green-up dates (d_{GU}) without further modification.

d_{GU} extraction is subject to a multitude of methodological and observational limitations (Helman, 2018) and involves considerably more – and often complex – steps after the original satellite measurement. These measurements are the seasonal progression of the Normalized Difference Vegetation Index (NDVI, Tucker (1979)) which exploits the vegetation's absorption characteristics in the red and near infrared spectrum of the sunlight and is a measure of photosynthetic active radiation absorbed by vegetation. In this context the NDVI evolution in temperate climates during spring time mainly represents different degrees of canopy closure. After raw satellite data retrieval, the methodological part starts with reconstructing the true NDVI profile, i.e. eliminating contaminated observations (due to clouds, haze etc.). Data from modern sensors such as MODIS (Moderate Resolution Imaging Spectrometer) or Sentinel-2 provide additional information such as quality measures but these are usually not available when analysing AVHRR (Advanced Very High Resolution Radiometer) time series. Fortunately, most nuisance factors lead to decreased NDVI observations, which (including knowledge about the phenological cycle in e.g. temperate climates) allows for constructing relatively simple filtering algorithms such as BISE (Best Index Slope Extraction, Viovy et al. (1992)), data compositing (Holben (1986)) or smoothing (Reed et al. (1994)).

Once contaminated observations are removed from the dataset the next step is to fit a function to the remaining NDVI observations. Here, interpolating and approximating methods are distinguished: linear interpolation (Viovy et al., 1992), spline interpolation (Hermance et al., 2007), logistic function (Fischer, 1994), Savitzky-Golay filter (Chen et al., 2004), Fourier transformation (Sellers et al., 1994) to name a few. Subsequently, phenological metrics are

computed based either on global or local thresholds (White et al. (1997); Duchemin et al. (1999)).

In summary, the process of extracting and matching phenological metrics from satellite and ground observations consists of five components: 1) Selection of satellite data, 2) Dynamic filtering/smoothing of raw NDVI values, 3) Fitting functions to the intra-annual NDVI time trajectory using selected NDVI observations, 4) Green-up date derivation (global or local thresholds), 5) Evaluation of satellite derived phenological metrics (potentially including comparison to ground observations).

So far research efforts have not yet led to a commonly accepted solution for this analytical procedure. As a result, each of these components can be (and often are) dealt with in a different way and the outcome – even of comparable study areas – do differ. This is an important issue because d_{GU} trends relying on different methods might leave us with ambiguous results and no indication as to what the phenological response to changing temperature and precipitation patterns is. In contrast to ground-based studies, consistent trends towards earlier start-of-the season have only been found when using a single method (Garonna et al. (2014) and references therein). When method ensembles were applied then i) no consistent trends could be extracted (White et al., 2009), ii) trends varied substantially dependent on land-cover types (Cong et al., 2013; Wang et al., 2015) or iii) even diverging trends between ground-based and remotely sensed data were found (Fu et al., 2014).

Most investigations on phenological trends have relied on the GIMMS dataset (Global Inventory Monitoring and Modeling System, Pinzon and Tucker (2014)) which supplies data from the 1980s onwards. Despite its inherent thorough satellite inter-calibration and correction for many other error sources this time series is limited due to its coarse temporal and spatial resolution. GIMMS represents only a subsample (averaged and dropped scan lines) of the highest possible spatial AVHRR resolution. AVHRR based products exist which e.g. offer daily observations at 1 km resolution over similar time periods. Such time series allow more robust and precise trend estimates for several reasons. First, daily imagery – rather than 15 day composites – capture the highly dynamic springtime phenological evolution at appropriate time scales as composite data might mask short term phenological changes (Fisher et al., 2006). Secondly, 1 km spatial resolution – rather than 8 km – ensures that many more observations stem from homogeneous land surface areas with just one land-cover type. The mixture of green-up signals received by the satellite hampers the extraction of phenological metrics for pixels with two or more underlying land-cover types especially in heterogeneous

3 *Satellite Phenology Products*

landscapes (Doktor et al., 2009). Trend estimates of such mixed pixels are subject to substantial uncertainties potentially intensified when different methods for metrics extraction are applied.

Unfortunately, almost all available NOAA AVHRR time series with this comparably high spatial and temporal resolution lack the necessary sensor inter-calibration and are also very difficult to acquire or cover a different geographic area (Eidenshink (2006)). In this study we had access to a unique time series covering the time span from 1989-2007 offering daily observations at 1 km resolution including sensor inter-calibration. This allowed us to investigate whether trend estimates from long-term satellite imagery of comparably high spatio-temporal resolution – where mixed effects should be less pronounced – confirm findings using the GIMMS dataset. Projects are underway to generate long and homogenised NOAA AVHRR time series globally (Global Land 1-km AVHRR Data Set Project) or over Europe (e.g. TIMELINE, DLR (2018)) but these either cover just a relatively short time period or have not been finished yet.

We hypothesise that varying trend estimates found between studies originate from mixed pixel effects captured differently by the methodological framework (s. five components above), especially with respect to the applied methods extracting phenological metrics. The outcome between these methods might vary dependent on land-cover type in general and the number of land-cover types within one satellite pixel. Ideally, trend estimates based on phenological metrics obtained with different extraction methods should exhibit only marginal differences. But due to their different mathematical nature some methods might relate to specific features within the NDVI profile which are not (or differently) captured by other methods connecting to developmental phenological stages. Nevertheless, these connections should be stable over time when satellite imagery of homogeneous land-cover is used, i.e. highest possible spatial resolution for forested areas. In conclusion, dependent on the underlying approach changes in the NDVI profile – related to different phenological timing, prevalent land-cover modification or crop rotation – might be captured differently despite observing the same phenomenon.

Consequently, this work is set-up as a twofold analysis. First, simulated satellite observations with implemented known trend to earlier green-up are generated in order to assess how a mixture of different green-up signals influences extracted metrics and respective trends. Here, the most commonly used methods to extract green-up dates are implemented. The seasonal NDVI progression originating from different land-cover types are simulated with

PROSAIL (Jacquemoud et al., 2009).

In a second step, the performance of these methods is tested across land-cover types based on real data. In addition to extracting phenological metrics and trends also a goodness-of-fit statistic across methods and land-cover types is employed. The use of NOAA AVHRR data – rather than e.g. MODIS data – is intended as the effects studied here are more pronounced in the former. d_{GU} of deciduous broadleaf forest are validated against respective d_{BB} from the extensive phenological network of the German Weather Service (DWD). Hence, the validation is performed upon a land-cover type with a stable geolocation. Additionally we employ a phenological model driven by temperature and photo period which has been shown to accurately predict d_{BB} (Schaber and Badeck, 2003; Lange et al., 2016).

3.2 Materials and Methods

The CORINE 2000 (COoRdinated INformation on the Environment) data base provides land-cover information at 250 m resolution (European Environment Agency, 2013). d_{GU} extraction (and trends thereof) was performed for the geographical area of Europe on land-cover classes which include natural and managed vegetation. The CORINE class 'broadleaf forest' was used for d_{GU} validation because i) of its stable geolocation over time, ii) due to respective ground data availability (Germany) and iii) the pronounced seasonal NDVI evolution. Each CORINE pixel of this class possibly consists of several deciduous tree species such as Beech or Oak. Furthermore, the NDVI pixels' reflectance is influenced by surrounding pixels which might exhibit different phenology.

3.2.1 Satellite data

The satellite data (NOAA 11, NOAA 14, NOAA 16, NOAA 17) were received and processed at the Meteorological Institute of the Freie Universität Berlin. The archive used here is the 'Mediterranean Extended Daily One Km AVHRR Data Set' (MEDOKADS, request via www.ufz.de/drp). MEDOKADS provides full resolution AVHRR data (all channels). This includes precise geolocation, cloud screening, correction for sensor degradation and orbital drift effects that cause non-linear changes in the signal measured, as well as inter-calibration between NOAA 11, NOAA 14 and NOAA 16 sensors to prevent inhomogeneities in the time

series (e.g. Koslowsky et al. (2001)). No correction of additional atmospheric effects is performed, neither any normalisation to account for the directional effects. North of 46° and east of 20° unstable transmission to the receiving ground segment caused occasional data gaps. Respective data was consequently removed from further analysis. More detailed information can be found in (Doktor et al., 2009).

The satellite pixels' land cover was classified on the basis of CORINE at 1 km resolution. In total 19 years (1989-2007) of daily high spatial resolution (1 km) NDVI data were analysed reflecting the seasonal phenological development. Unfortunately, MEDOKADS was discontinued from the year 2008 onwards. The authors had no access to comparable, i.e. daily 1 km resolution, AVHRR datasets for the years 2008 and beyond. Therefore, the analysis had to be restricted to the time period mentioned above. The MEDOKADS dataset has predominantly been used for the analysis of land-cover change and desertification (e.g. Hill et al. (2008); del Barrio et al. (2010); Stellmes et al. (2013)). More detailed information regarding data processing can be found in Friedrich and Koslowsky (2009).

NDVI values based on daily Terra MODIS (MODerate resolution Imaging Spectrometer) reflectances (MOD09GQ, Version 5, 231 m spatial resolution) were used to assess potentially undetected drifts within the MEDOKADS dataset. Data can be accessed via GloVis, USGS Earth Explorer or via ftp download: [e4ftl01.cr.usgs.gov](ftp://e4ftl01.cr.usgs.gov)

3.2.2 Ground observational data

Yearly extracted d_{GU} were evaluated against respective d_{BB} as collected by the extensive observational DWD network. The phenological records cover a range of deciduous tree species and provide about 1800 observations per year and species respectively which are evenly distributed over Germany (Fig. 3.1). We used the mean d_{BB} of the dominating tree species Beech and Oak to compare with the mean d_{GU} of NDVI satellite pixels classified as deciduous forest. Phenological observation stations do not necessarily geographically coincide with the CORINE land-cover class of broadleaf forest. They do however, exhibit the seasonal phenological evolution of a species at the landscape to nation-wide level. The satellite spatially integrates anyway such that single plant observations and geographically coinciding NDVI-values often differ substantially (Badeck et al., 2004) which can only be accounted for by intense and regular monitoring on the ground (Liang et al., 2011). The phenological data was extracted out of the phenological database PPODB (Dierenbach et al., 2013) based

on respective databases of the German Weather Service.

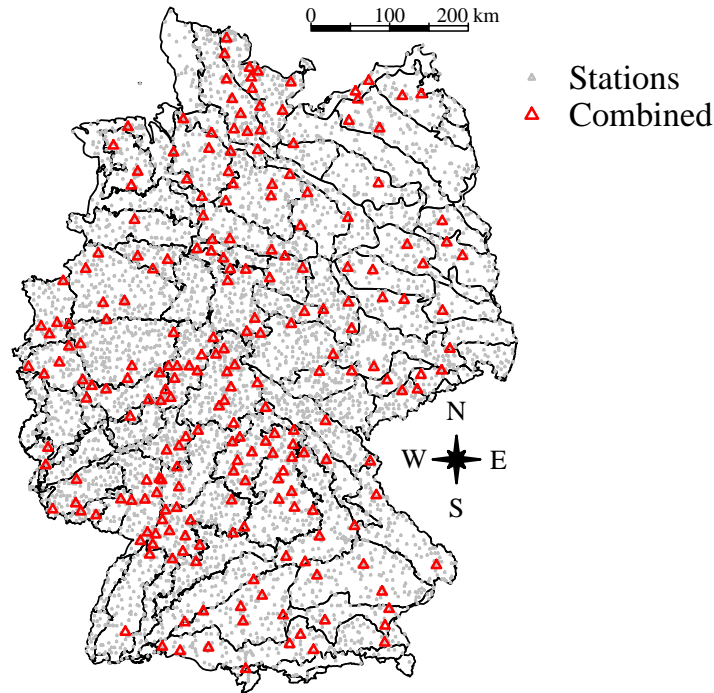


Figure 3.1: Phenological observation stations (grey triangles) within Germany which were active between the years 1951 to 2009. For each natural region (black lines, a region which is considered to be homogeneous with respect to topography, soil, climate and vegetation) at least one (or more) combined time series was computed (red triangles) and subsequently used for phenological modelling.

Phenological data are often fragmented including observation gaps and/or time series of different lengths at different observation stations. For calculating trends of phenological ground observations we therefore combined all available time series using the R-package *pheno* (Schaber, 2012). Fragmented time series can be combined to a continuous time series, when the observing stations are geographically close enough and if their time series overlap (Schaber and Badeck, 2002). During that process outliers are also eliminated. Stations with a maximum distance of ten kilometers and a maximum difference in altitude of 50 meters are combined. Only station clusters with a minimal size of five stations are used to ensure the reliability of the resulting time series. The resulting combined time series is much closer to the land surface phenology sensed by the satellite than that of a single station.

3.2.3 Bud burst modelling

The Promoter-Inhibitor Model (PIM) was introduced by Schaber and Badeck (2003) simulating bud burst of deciduous trees based on interactions between hormones or compounds controlling a plant's life-cycle. The underlying assumption is that the ratio between these agents controls the different growth phases and that these agents are primarily controlled by temperature and photo-period. We employed this model to simulate d_{BB} of tree species (s. also Lange et al. (2016)) using the R-package *phenmod* (Lange, 2013). For a detailed description of the model please refer to Schaber and Badeck (2003).

3.2.4 Simulated NDVI time series

The methods' performance extracting phenological metrics was first tested on simulated NDVI time series stemming from 1) homogeneous land-cover, i.e. a single underlying land-cover type within a satellite pixel and 2) heterogeneous land-cover, i.e. two or more different underlying land-cover types within a satellite pixel. Regarding NOAA AVHRR's spatial resolution of 1 km case 2) is much more common in the heterogeneous landscapes of Europe. The NDVI's seasonal evolution for many land-use / land-cover types usually follows a logistic (or double logistic) function (Fischer (1994)):

$$NDVI(t) = vb + \frac{k}{1 + e^{-c(t-p)}} - \frac{k + vb - ve}{1 + e^{-d(t-q)}} \quad (3.1)$$

where k = the asymptotic value, c = the slope of the first inflection point, d = the slope second inflection point, p = the date of first inflection point, q = the date of second inflection point, vb = the NDVI value before green-up, ve = the NDVI value at the end of senescence and t = the time variable (Day of Year, 1 - 365). Many other studies also found that the NDVI seasonal evolution follows a logistic shape across land-cover types using satellite data or temporally dense spectral ground measurements (e.g. Huemmrich et al. (1999); Wittich and Kraft (2008); Zhang et al. (2006); Soudani et al. (2012); Nestola et al. (2016)).

The non-linear mixing nature of the NDVI signal (i.e. ratio of the sum/difference of spectral channels) requires a simulation of surface reflectances in the red and near-infrared part of the electromagnetic spectrum (Lobell and Asner, 2004; Jiang et al., 2006). We simulated daily surface reflectances over the course of a calendar year using the radiative transfer model

PROSAIL (Jacquemoud et al. (2009)) for four land-cover types: winter crops, summer crops, pastures and deciduous forests. Relevant parameters were modified according to their seasonal evolution in temperate climates (s. Equation 3.1 and Figure 3.2).

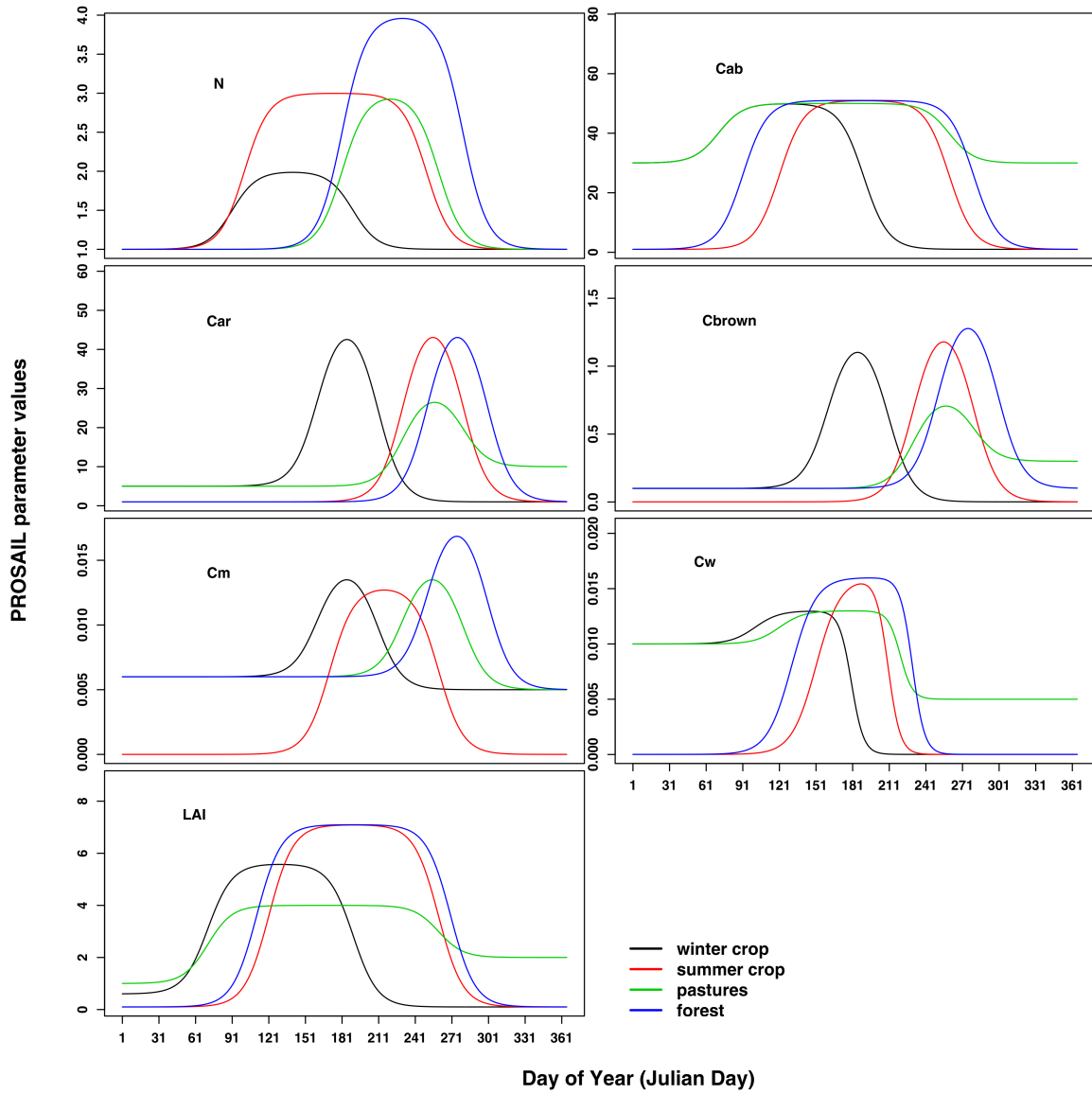


Figure 3.2: Implemented seasonal evolution of the PROSAIL parameters C_{ab} , C_{ar} , C_{brown} , C_w , C_m and N . PROSAIL simulations were performed for four different land-cover types: winter crops (black line), summer crops (red line), pastures (green line) and forests (blue line).

These parameters included Chlorophyll (C_{ab}), Carotenoids (C_{ar}), Equivalent water thickness

(C_w), Dry matter content (C_m) and Leaf Area Index (LAI). Varying parameter values with slightly different timings would have been plausible too. But here, the primary objective was to set-up a theoretical scheme which allows to identify the performance and applicability of methods to extract phenological metrics stemming from homogeneous versus heterogeneous underlying land-cover. The resulting spectral profiles at a resolution of 1 nm were spectrally resampled to generate AVHRR band 1 (0.58 - 0.68 μm) and band 2 (0.725 - 1.00 μm). Simulations were set-up exemplary for a latitude of 50° North with a corresponding solar zenith angle for every Day of the Year (DOY), an observer zenith of 10°, a relative azimuth of 140° with a leaf angle distribution representative for erectophiles (crops).

Mixture type	Land-covers	Area proportions
Equal area	Winter crop, Summer crop, Pastures	33.3% each LC
Uneven area	Winter crop, Summer crop	80% vs 20%
Arbitrary area	Winter crop, Summer crop, Pastures	between 10% - 50%
Spatial trend	Winter crop, Summer crop, Pastures	Decrease/increase ca. 1.5% per year

Table 3.1: The four land-cover mixture type cases for which NDVI values were simulated with PRO-SAIL: Equal area, uneven area, arbitrary area and spatial trend.

From these daily AVHRR band 1 and 2 simulated reflectances we generated 19 seasonal NDVI profiles (in line with the available real data from 1989-2007) for every land-cover type. An artificial underlying trend towards earlier green-up (1 day/year) was introduced by changing p (date of the first inflection point) for all relevant PROSAIL parameters. Additionally, AVHRR sensor noise was added on all simulated NDVI values according to:

$$NDVI_{i*} = NDVI_i + \mathcal{N}(\mu, \sigma^2) \quad (3.2)$$

where $\mu = 0$ and $\sigma = 0.02 + 0.02 * NDVI_i$ (Terra MODIS sensor noise, Vermote and Kotchenova (2008a)). Out of these daily simulations NDVI values were selected according to the likelihood of clear sky observations as found by Hmimina et al. (2013) for further analysis. Thus, introducing characteristics of real satellite data time series, where many NDVI values in temperate climates cannot be used due to frequent cloud cover or haze and the likelihood of clear sky observations increases in summer time. This selection procedure was repeated 100 times in order to derive stable trend estimates. The resulting homogeneous seasonal NDVI profiles were spectrally mixed according to predefined area ratios within one

satellite pixel simulating the NDVI signal of a heterogeneous land-cover type within one satellite pixel. Here, we assumed four cases (*Equal area*, *Uneven area*, *Arbitrary area* and *Spatial trend*) with 2-3 land-cover types respectively (s. Table 3.1).

These cases shall represent typical landscape configurations encountered in Central and Western Europe. *Equal* and *Uneven area* assume a temporally constant mixture of identical land-covers over time with different area proportions each. The *Arbitrary area* case assumes i) yearly crop rotation and ii) also inherent slightly varying geolocation of NOAA AVHRR imagery. Both phenomena combined can lead to substantially changing area proportions of 3 land-cover types from year to year. The *Spatial trend* case considers a human induced gradual and linear underlying land-cover change over time.

Simulated seasonal NDVI profiles vary considerably between homogeneous land-cover and also between the four cases (s. Figure 3.3). Subsequently, d_{GU} and trends thereof were extracted for all simulated NDVI seasonal profiles. Methods for d_{GU} extraction are presented in the following section.

3.2.5 Extracting phenological metrics

Daily observations were dynamically filtered using an adaptation of the Best Index Slope Extraction (BISE) algorithm (Viovy et al. (1992)). The phenological cycle's seasonal evolution in temperate climates is well known and used to detect and eliminate cloudy observations. For daily NDVI observations over the course of one year, the algorithm only accepts decreasing NDVI values if no higher value is found within a so called sliding period (window). A period of 40 days showed best results for our study area. Study areas of different climates, e.g. tropics, might require a modified (longer) sliding period. Simulated NDVI values were not filtered.

After dynamic filtering, functions were fitted to the remaining NDVI observations using a variety of methods (s. Table 3.2). Here, we referred to differential evolution techniques using the R-package *DEoptim* (Ardia et al., 2011) for optimisation. Subsequently, the influence of varying thresholds on respective d_{GU} was tested: A set of three dynamic (local) and two global thresholds were applied on the modelled NDVI curve. Local thresholds (LT) are defined by the intra-annually extracted NDVI maximum and minimum values of a single pixel:

$$LT = NDVI_{min} + (NDVI_{max} - NDVI_{min}) * k \quad (3.3)$$

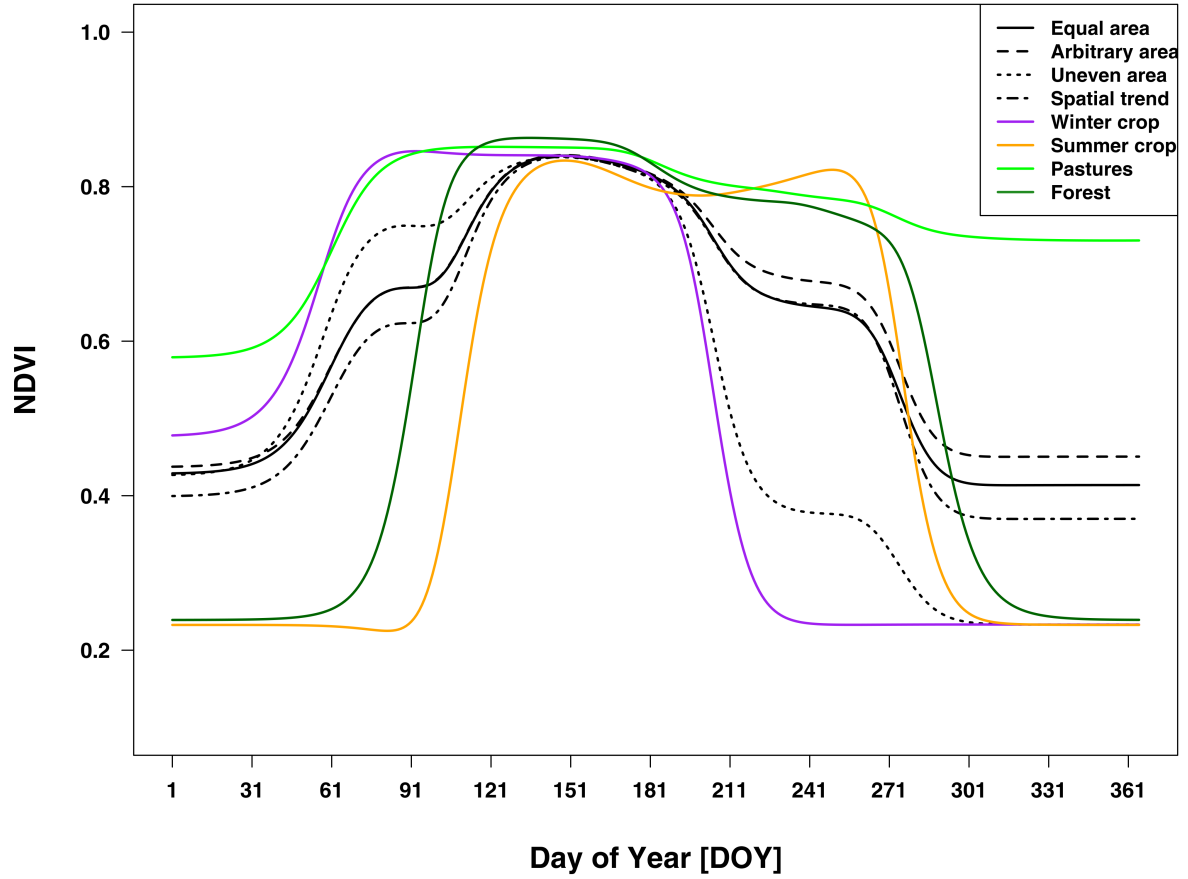


Figure 3.3: Simulated seasonal NDVI profiles of an exemplary year considering homogeneous land-cover including winter crops (purple line), summer crops (orange line), pastures (green line) and deciduous forest (darkgreen line) as well as the four cases of heterogeneous land-cover *Equal area* (solid black), *Arbitrary area* (long dashed), *Uneven area* (short dashed) and *Spatial trend* (short/long dashed).

We defined k at 0.55, 0.60 and 0.65 ($kl55$, $kl60$ and $kl65$). That means that LT can refer to different NDVI values for the same pixel between years and also to different NDVI values for neighbouring pixels within one year. Global thresholds (GT) on the other hand refer to fix NDVI values of 0.35 and 0.40 ($kg35$ and $kg40$) for any year and pixel. The thresholds values were chosen empirically against observed bud burst observations of the German Weather Service.

Method	abbreviation	reference
Linear Interpolation	<i>Li</i>	R-func. 'approxfun'
Monotone Hermite Spline	<i>Sp</i>	R-funct. 'splinefun'
Sigmoidal function	<i>Sg</i>	R implementation (Zhang et al., 2003)
Logistic function	<i>Lg</i>	R implementation (Fischer, 1994)
Gaussian function	<i>Gs</i>	C implementation (Press, 1992)
Savitzky-Golay filter	<i>SG</i>	C implementation (Press, 1992)
Fast-Fourier Transformation	<i>FFT</i>	R-func. 'fft'

Table 3.2: Methods used for extracting phenological metrics from NDVI observations. All methods are implemented in 'R' either using R-functions or embedded C-code and are freely available through the R-package 'phenex'. An initial dynamic filtering (BISE) is applied on all satellite based daily NDVI data beforehand.

3.2.6 Trend analysis

For trend estimates of d_{GU} time series (1989-2007, for each pixel and method) we used a linear regression model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ where Y is the dependent variable, β_0 the intercept, β_1 the slope, X the explanatory variable and ε the error. Here, d_{GU} is the dependent and the observation year the explanatory variable ($d_{GU} \sim obsYear$). First, a test regarding temporal autocovariance or autocorrelation was performed (the residuals of modelled relationships were checked using Local Indicators of Spatial Association, Anselin (1995)) for a subset of d_{GU} time series (all broadleaf forest pixels within the area of Germany for all years analysed). As no or only slight autocorrelations were found, there was no need to refer to autoregressive time series models and the linear regression model shown above was applied. Spearman's correlation was not used for trend analysis as this method only indicates positive or negative trends but gives no indication on the trend magnitude, i.e. days/year. Furthermore, significance levels differ between Spearman's correlation and linear regression. The former tends to extract higher significance levels, i.e. lower absolute values thereof, than the latter.

3.2.7 Goodness-of-fit

Goodness-of-fit (R^2 , coefficient of determination) is – next to the comparison to d_{BB} – another criterion to assess a methods applicability for phenological metrics extraction. R^2 was

derived from modelled versus observed (BISE corrected NDVI observations) values for all approximating methods of satellite based data.

3.3 Results

3.3.1 Satellite vs ground observations

d_{GU} estimates for broadleaf forests in Germany are on average substantially closer to d_{BB} using local rather than global thresholds (Table 3.3). For *kl65* most methods perform best with mean absolute deviations between 6-8 days ($Lg \sim 15$ days). d_{GU} using *kg35* or *kg40* differ ~ 2 -3 weeks. Most likely another set with increased thresholds, e.g. *kl70* or *kg45*, would further reduce these deviations. Due to time constraints this was not tested here. But especially higher global thresholds could cause difficulties to apply the implemented algorithms (i.e. the NDVI threshold is close to or as high as the NDVI maximum value). In the following only results for *kl65* are presented as this dynamic threshold performed best when compared to ground observations.

Method	<i>kl55</i>	<i>kl60</i>	<i>kl65</i>	<i>kg35</i>	<i>kg40</i>
Daily observations					
<i>Li</i>	17.9	12.0	7.5	24.9	13.9
<i>Sp</i>	15.0	9.5	6.7	25.2	14.0
<i>FFT</i>	16.8	11.2	7.2	27.3	15.5
<i>SG</i>	13.0	9.2	6.0	18.4	10.6
<i>Sg</i>	15.0	9.5	6.7	25.2	14.0
<i>Gs</i>	18.8	13.3	8.5	27.4	16.1
<i>Lg</i>	22.4	18.7	15.4	26.3	18.2

Table 3.3: Mean absolute deviation of computed green-up dates (d_{GU}) across methods (rows) and thresholds (columns) for broadleaf forests in Germany compared to respective ground observations (d_{BB}).

3.3.2 Trends from simulated NDVI profiles

Detected trend magnitudes for homogeneous land-use (1 land-use per satellite pixel) on the basis of daily NDVI values are in line with the actual implemented trend (1 day/year) for

the interpolating methods (s. Table 3.4). Most approximating methods with the exception of *Lg* detected slightly varying trends. In case of simulated mixed pixels, i.e. two or more different land-use types, differences between actual and detected trends are substantial (up to 2 days/year). Differences are most noticeable for approximating methods and the *Spatial trend* case.

	Dec. forest	Summer crop	Pastures	Winter crop	Spatial trend	Uneven area	Equal area	Arbitrary area
<i>Li</i>	-0.88	-0.95	-1.09	-0.61	0.35	-0.80	-0.86	-0.19
<i>Sp</i>	-0.87	-0.98	-1.09	-0.65	0.48	-0.70	-0.81	-0.15
<i>Sg</i>	-0.81	-0.81	-0.69	-0.69	1.15	-0.80	-0.98	-0.19
<i>Lg</i>	-0.85	-0.76	-1.13	-0.84	0.13	-0.68	-0.55	-0.18
<i>Gs</i>	-0.60	-0.65	-0.35	-0.74	0.29	-1.43	-0.70	-0.37
<i>FFT</i>	-0.85	-0.70	-0.60	-0.32	-0.41	-0.51	-0.66	-0.20
<i>SG</i>	-0.61	-0.57	-0.57	-0.83	-0.17	-0.69	-0.83	-0.31

Table 3.4: Trends (days/year) for simulated NDVI time series with daily resolution excluding sensor noise.

Results differ if simulated NDVI values are only considered for trend analysis according to clear sky probability (s. Figure 3.4). Here, trend estimates from approximating methods do differ less from interpolating methods, especially method *Lg*. Still, the detected trend differences are on average between 0.5 - 1 days/year, most noticeable for the cases *Even area* and *Uneven area* where mostly positive trends – the highest from interpolating methods – were extracted from simulated data. All methods used in theoretical framework underestimate the simulated trend.

3.3.3 Goodness-of-fit

R^2 varies between approximating methods and across land-use classes for the area of Europe (Figure 3.5). Interpolating methods obviously exhibit a perfect fit between model and BISE selected NDVI observations and are therefore not shown here. *Lg* shows high R^2 values across all land-use classes whereas all other methods vary more between but also within land-use classes, especially *Sg* and *Gs* thereby confirming results of the previous section, where differences between simulated and extracted trends were also pronounced for these

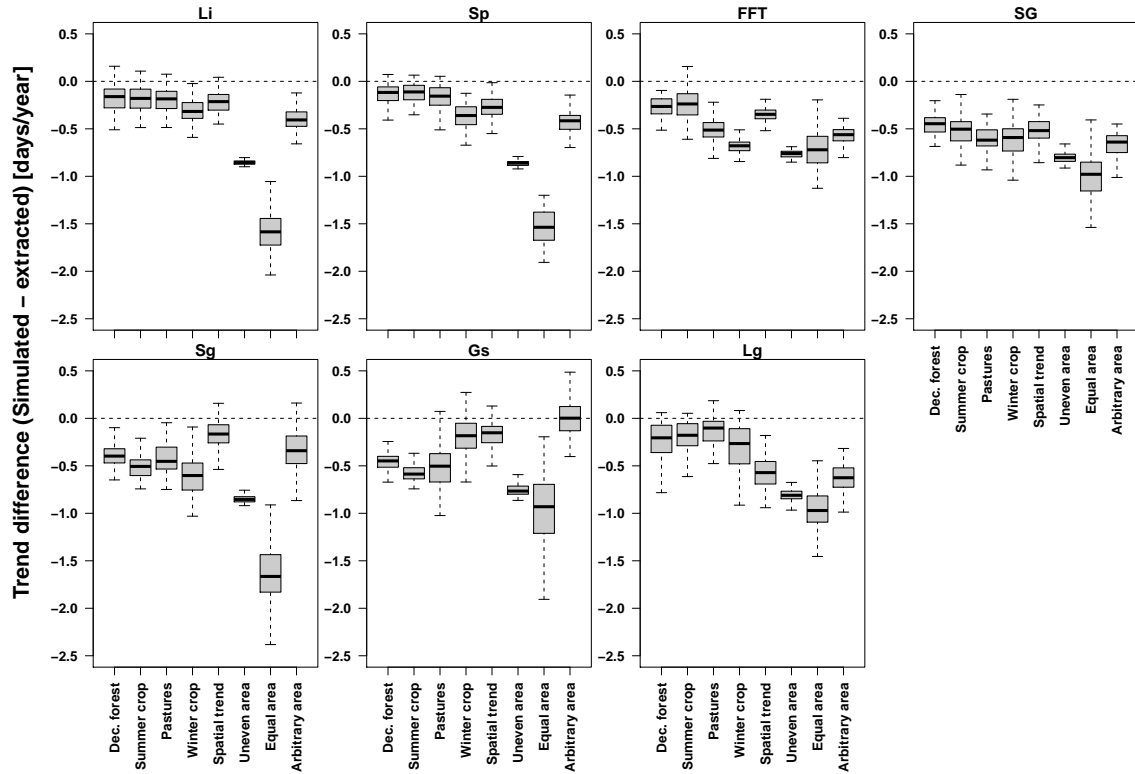


Figure 3.4: Trend differences (simulated - extracted) [days/year] across methods for i) a single land-cover within a satellite pixel (Dec. forest, Summer crop, Pastures, Winter crop) and ii) mixed land-cover within a satellite pixel (*Spatial trend*, *Uneven area*, *Equal area*, *Arbitrary area*). Trend estimates are based on selected simulated NDVI values (for 19 years, 100 repetitions) – according to the likelihood of clear sky observations – with an induced trend of 1 day/year towards earlier green-up.

methods. Land-cover classes exhibiting low R^2 values are: Agro-forestry, Olive groves, fruit plantations, and Sclerophyllous vegetation.

Best fits are achieved for the land-cover classes 'non-irrigated arable', 'pastures', 'rice fields' and the different forest classes which combined constitute $\sim 75\%$ of the analysed area.

3.3.4 Green-up dates across methods

All methods produce on average a Southwest-Northeast gradient towards later green-up across Europe (Figure 3.6 and 3.7). This pattern goes in line with the prevailing European climatic conditions. A substantial part of areas with $d_{GU} < 50$ is covered by evergreen vegeta-

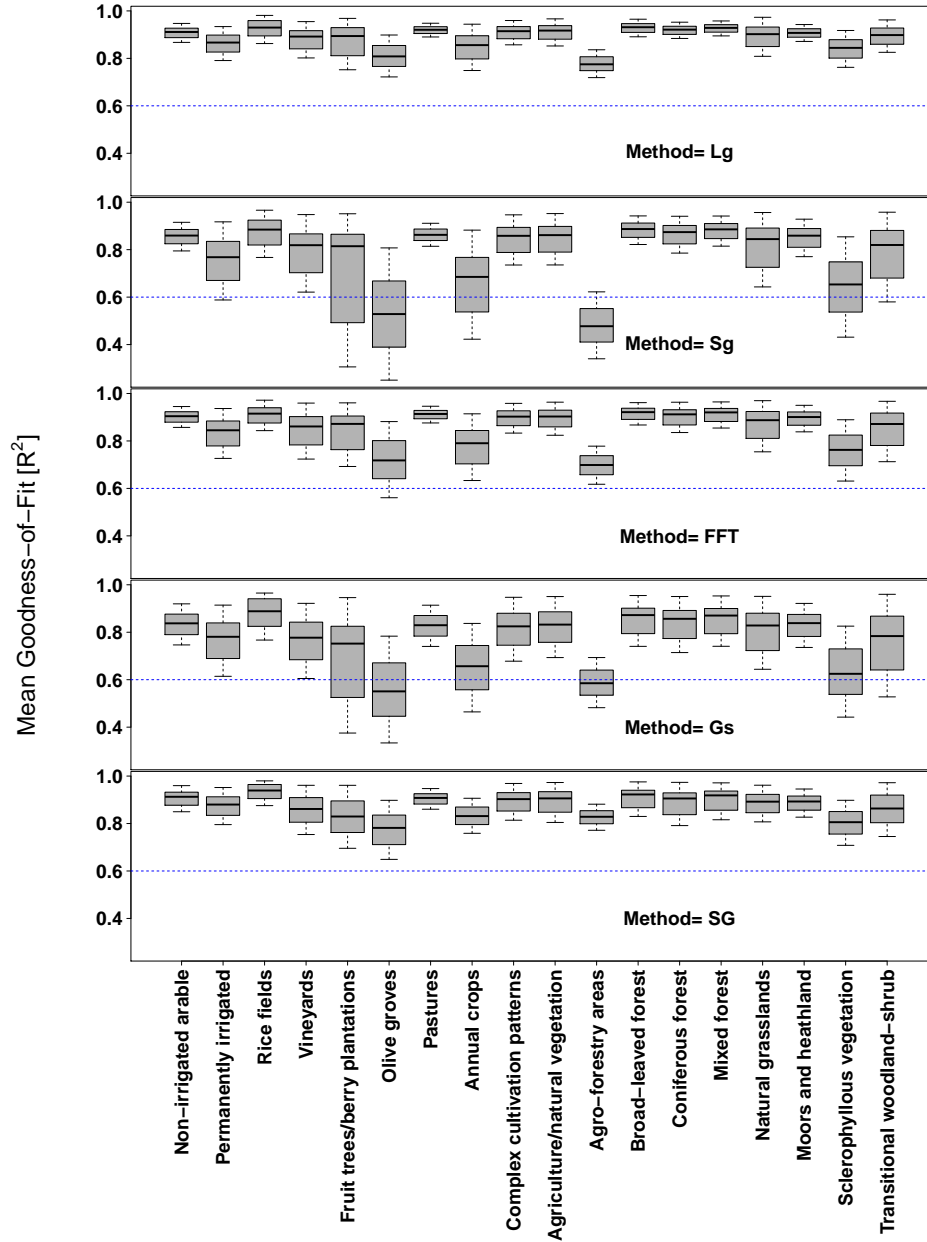


Figure 3.5: Average goodness-of-fit, R^2 , for CORINE land-use classes (x-axis) and approximating methods (*Lg*, *Sg*, *FFT*, *Gs*, *SG*), i.e. excluding *Li*, *Sp*. R^2 values are based on the period 1989-2007 for fits on daily NDVI observations. The dashed blue line indicates R^2 values of 0.6.

tion (southern Europe) or pastures (France and England). Mountainous regions like the Alps or Pyrenees exhibit a later d_{GU} (> 140) than lowlands like the Loire, Rhine or Po valley ($<$

3 Satellite Phenology Products

110). Some agricultural landscapes in those areas are subject to irrigation or even controlled flooding such as rice fields in northern Italy. Here, the satellite cannot receive vegetation signals due to surfaces covered extensively by water one month before harvest. This consequently results in comparably late d_{GU} . Mean d_{GU} vary spatially across methods (s. Figure 3.7) up to 50 days for a given pixel.

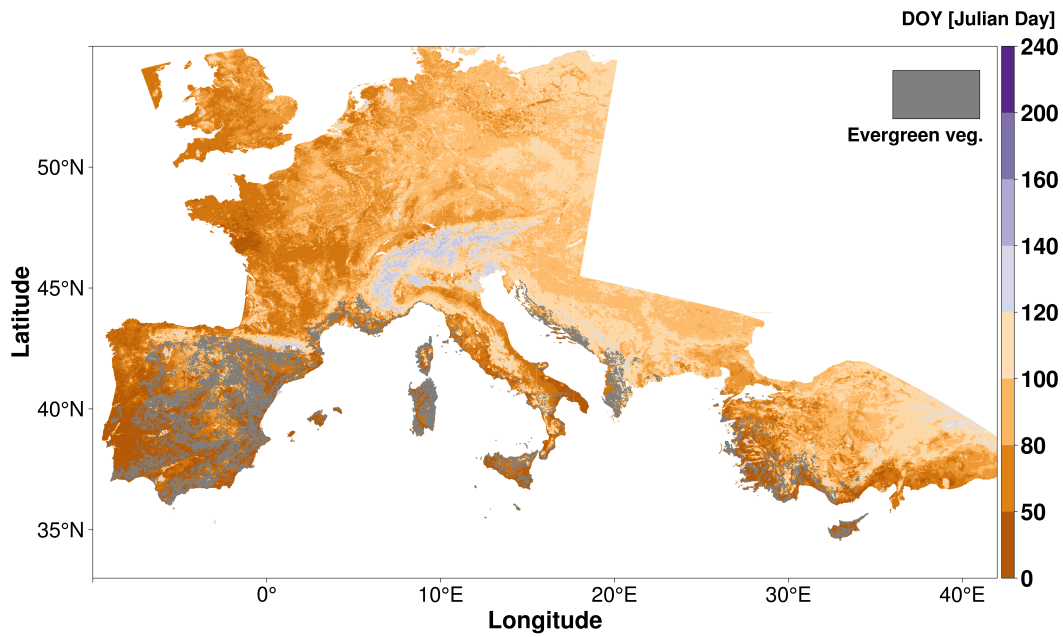


Figure 3.6: Mean d_{GU} of the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold ($kl65$) using linear interpolation (Li).

Phenological metrics based on Sp vary slightly from Li between 0 - 10 days mean difference towards later d_{GU}). This picture of a spatially relatively coherent difference to Li changes

for the remaining methods. *Lg* extracts on average earlier d_{GU} (0-10 days) for agricultural areas and conifers but later dates for mountainous areas and pastures. *Sg* extracts later d_{GU} for pastures and forests but earlier d_{GU} for crops which is mostly in line with *Gs*. Similar spatial patterns are found for *FFT* and *SG* but on a generally lower magnitude of differences compared to *Li*.

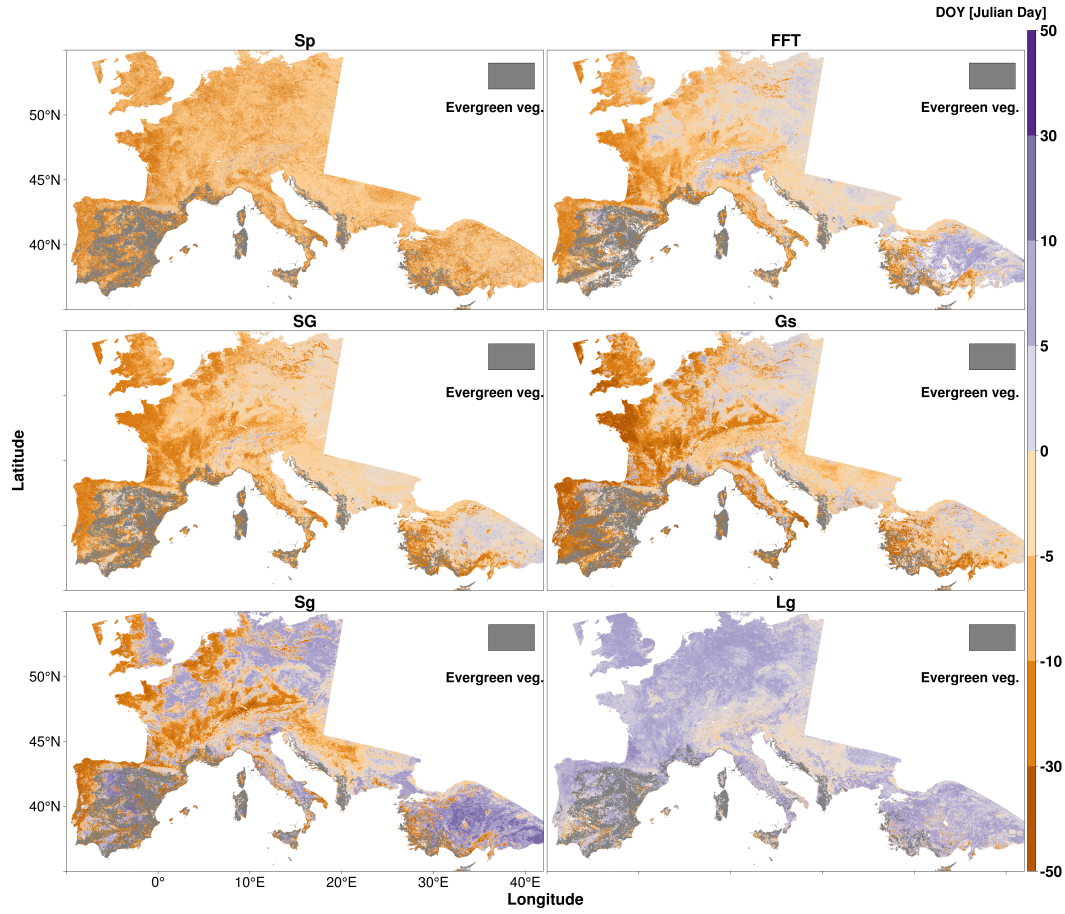


Figure 3.7: Mean d_{GU} of the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold (*kl65*) for all other methods. All plots show the difference of the respective method to *Li* (i.e. Figure 3.6).

3.3.5 Trends from satellite observations

Trend estimates for many land-cover classes are subject to potentially disturbing factors such as underlying land-use change (unstable geolocation) or crop rotation. We therefore firstly analysed specifically trends derived for the class of broadleaf forests in Germany because of the inherent stable geolocation and the possibility to compare trends of satellite and ground observations (Figure 3.8, grey and green boxplots). Trend estimates based on modelling are also integrated into the analysis (red boxplot).

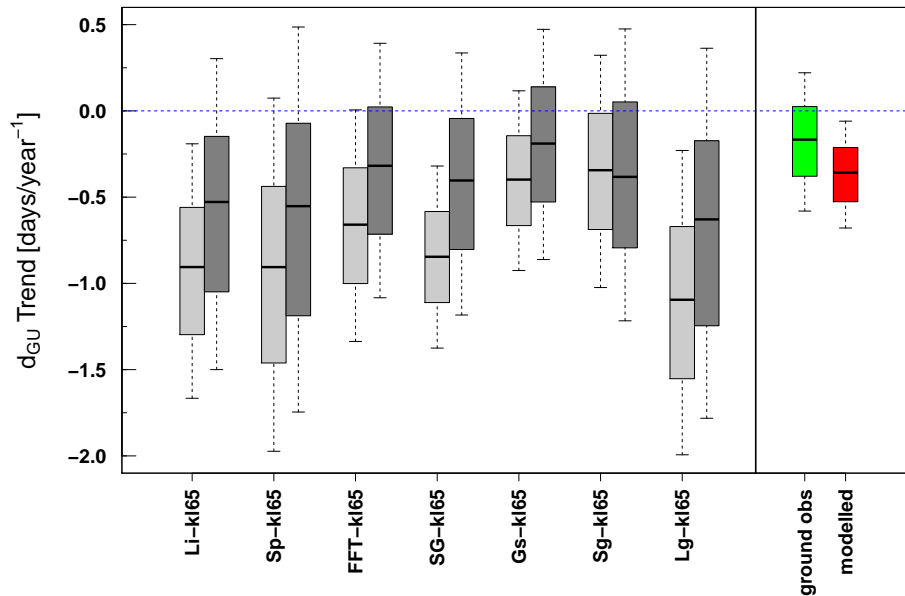


Figure 3.8: Trend [days/year] of all tested methods for the period 1989-2007 based on daily NDVI observations for areas of deciduous forest in Germany (light grey) and Europe (dark grey) using a dynamic (local, $kl65$) threshold. Trend estimates for ground observations are shown in green and for modelled phenology (both Germany) in red boxplots. The dashed blue line divides positive and negative trend estimates.

Trend estimates based on ground observations and modelling show a moderate spring advancement for broadleaf forests concerning the period from 1989-2007 of ~ -0.3 days/year. Dynamic threshold (kl) based d_{GU} trends also indicate a moderate spring advancement (0.6 - 1.1 days/year on average) across methods and the area of Germany (light grey boxplots).

Within this homogeneous pattern of earlier d_{GU} , differences between methods are still evident and can be discriminated into 3 groups: Interpolating methods (Li , Sp), filtering methods (SG , FFT) and approximating methods (Sg , Gs , excluding Lg). The mean trend of approximating methods differs about ~ -0.4 days/year compared to ground references. Filtering and interpolating methods, however, indicate trend differences of ~ -0.8 days/year and Lg even up ~ -1.0 days/year. This translates over the observed time period into 2-3 weeks deviation. The trend inter-quantile ranges also vary between methods, e.g. SG versus Sp .

In a second step the analysis was extended including all deciduous forests in Central and Western Europe (dark grey bars Figure 3.8). Here, results largely mirror those produced for the smaller scale (Germany), concerning relative differences between methods. The absolute trend magnitudes though (towards spring advancement) are less pronounced.

3 Satellite Phenology Products

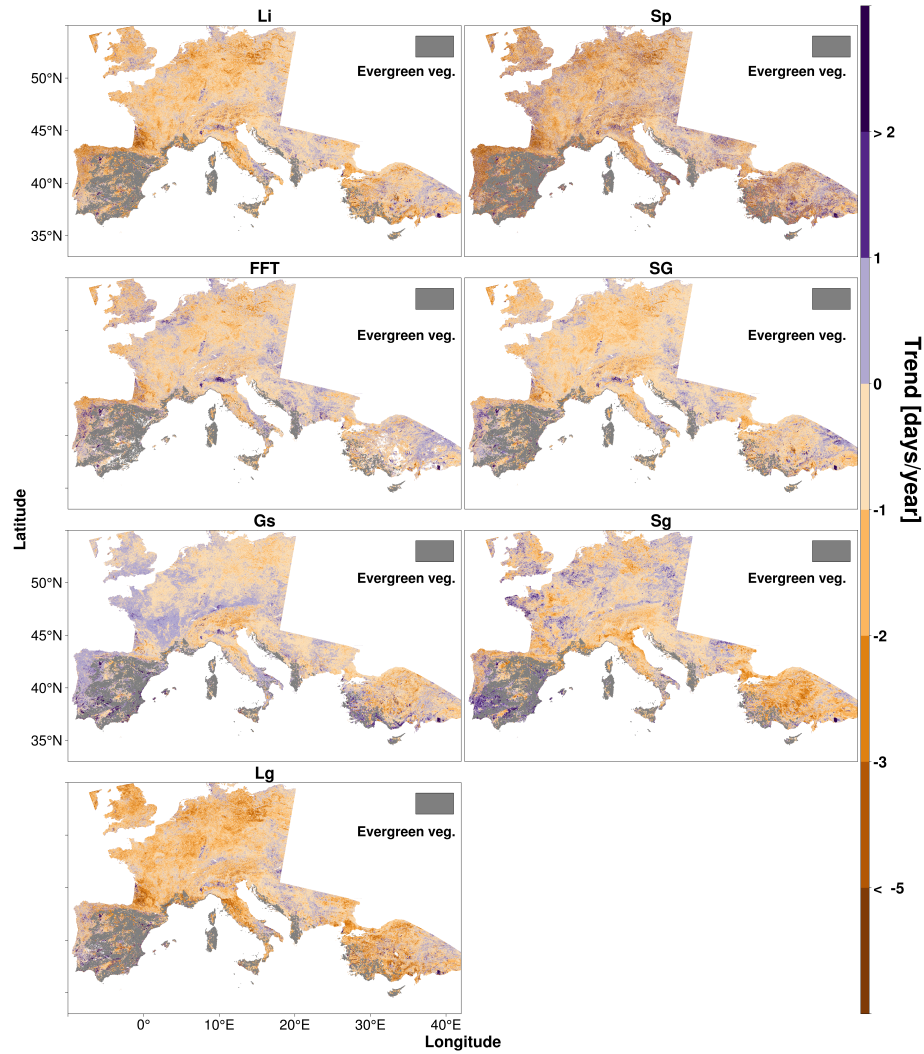


Figure 3.9: Trend [days/year] of the period 1989-2007 for Europe and the Mediterranean based on daily NDVI observations and a local threshold (*kl65*) for all methods: *Li*, *Sp*, *FFT*, *SG*, *Gs*, *Sg* and *Lg*.

Mapped trend estimates across Europe confirm the above and also show distinct land-cover dependent spatial heterogeneities which vary dependent on the underlying method (Figure 3.9). The majority of trend estimates across and land-cover classes and methods is between -2 and 0 days/year towards spring advancement (between 60-70%, s. Table 3.5). Stronger positive or negative trends are mostly occurring in agricultural areas (mostly river valleys in Germany, Italy, France and Spain) indicating no spring advancement but rather a change in cultivated crops.

Dartmoor, a moor in southern England, shows a particularly strong negative trend, pointing at the vulnerability of these ecosystem types. Positive trends in many grasslands might point to an increased intensification (fertilisation) which results in higher NDVI values artificially delaying derived d_{GU} dates.

Methods/Trend	< -5	-5 - -2	-2 - -1	-1 - 0	0 - 1	1 - 2	> 2	[d/y]
Li	0.6%	10.2%	25.7%	46.2%	14.8%	1.8%	0.7%	
Sp	2.0%	14.7%	21.5%	34.7%	17.0%	5.5%	4.5%	
FFT	0.1%	4.3%	17.7%	50.1%	23.8%	3.1%	0.8%	
SG	0.1%	3.3%	20.9%	55.2%	17.5%	2.1%	0.8%	
Gs	0.1%	2.7%	9.5%	50.6%	33.4%	2.9%	1.3%	
Sg	0.1%	5.9%	21.1%	45.2%	21.8%	5.0%	1.1%	
Lg	0.7%	13.7%	29.9%	39.7%	12.7%	1.7%	1.6%	

Table 3.5: Percentage of pixels per method falling into discrete trend categories (d/y, days per year) also used for mapping (s. Figure 3.9).

Trend estimates vary between the major CORINE land-use classes (European scale, Figure 3.10). Despite an on average moderate d_{GU} advancement, the mean trend difference between land-cover classes is up to 1 day/year. Some land-cover types even exhibit method dependent positive trends, e.g. agro-forestry areas, rice fields or annual crops. The inter-quantile range is generally lowest for broadleaf forest and highest for agro-forest or annual crops. The latter are also subject to maximal difference of inter-quantile ranges between methods.

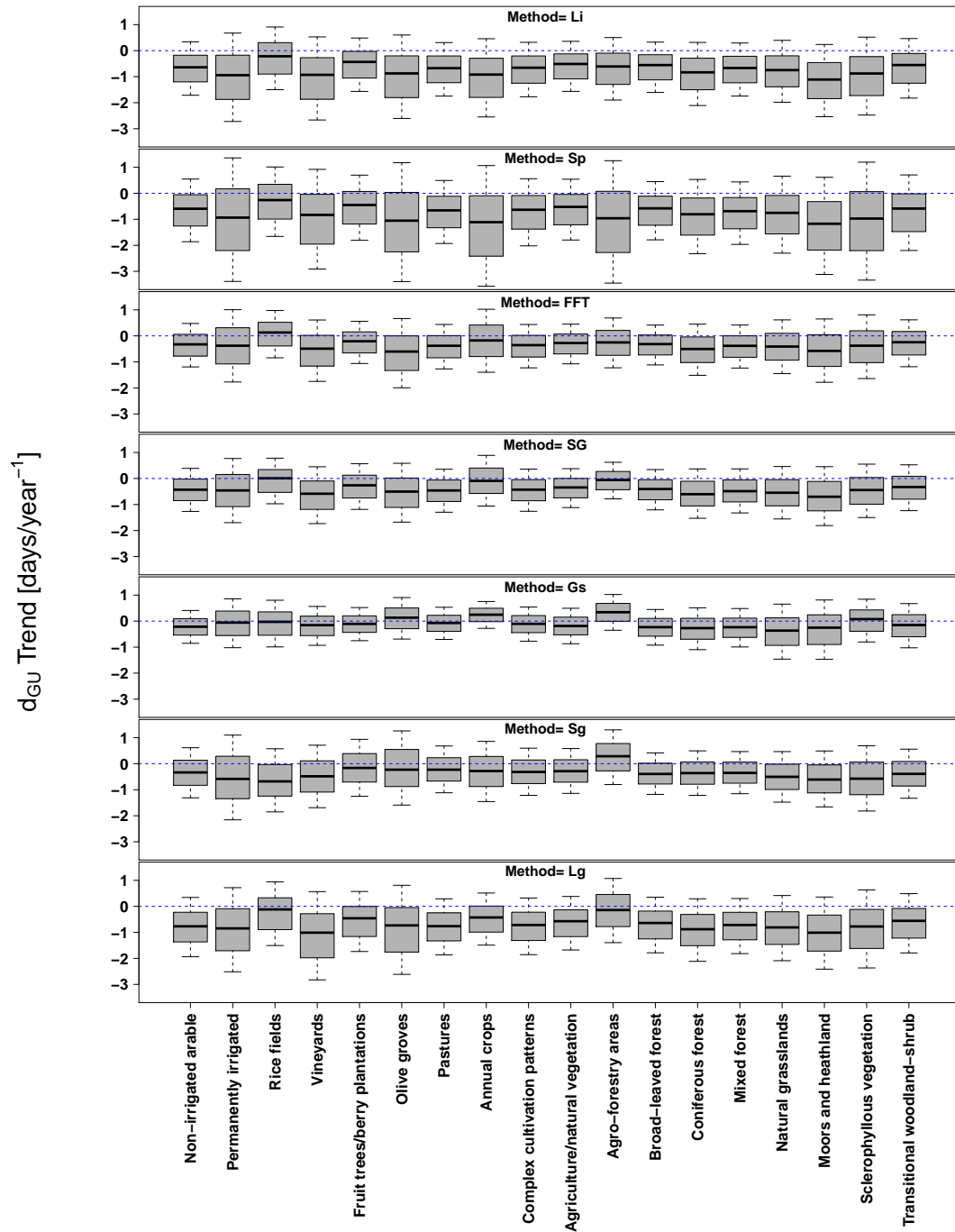


Figure 3.10: d_{GU} trend (days/year) over the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold ($kl65$) across major CORINE land-use classes.

3.4 Discussion and conclusions

3.4.1 Trends from simulated NDVI time series

In case of homogeneous land-use, the extracted trend differences between land-cover / land-use classes are most likely due to varying seasonal NDVI amplitudes. A shifting NDVI signal (towards earlier green-up) is captured less well by functions resembling a normal distribution (e.g. G_s , FFT , S_g) especially if the NDVI amplitude between winter and late spring is less pronounced, as in winter crops and grasslands. For heterogeneous landscapes, the underlying trend is blurred by diverging green-up signals altering the shape of the NDVI' seasonal evolution. Interpolating methods are more suitable to extract trends in mixed NDVI signals (with exception of L_g), as this signal differs from the normal distribution (shoulders left and right of NDVI maximum). Hence, respective methods (G_s , FFT , S_g) fit less well onto NDVI profiles of single years. Trend estimates are consequently associated with higher uncertainties.

Substantial differences between approximating and interpolating methods decrease on simulated NDVI time series with a lower temporal resolution. This is to be expected as the shoulders left and right of the NDVI maximum are not as distinct anymore, resembling more a normal distribution. This comes along with decreasing performance of interpolating methods. For more details on the influence of NDVI data gaps on phenological metrics extraction please refer to Hmimina et al. (2013).

Considerable differences of phenological metrics using high and medium resolution satellites have been found where land-use is unevenly represented within a remotely sensed pixel (Zhang et al., 2017). This is in line with our case *Uneven area* where extracted trend estimates deviate substantially from the true underlying trend. Analysing the seasonal NDVI profiles and associated d_{GU} for, e.g. the case *Spatial Trend*, reveals that these deviating trends are often driven by the smaller land-use fraction and/or the land-use with the later phenological signal (here summer crop/maize). Chen et al. (2018) found a varying sensitivity between tested methods to changes in the growing season based on simulated data, confirming our results.

Assigning CORINE land-cover classes to the theoretical framework of this study is not a trivial task. Due to NOAA AVHRR's coarse spatial resolution almost all land-cover classes are potentially subject to mixed pixels and land-use change processes. Nevertheless, it is to

be expected that the classes *Permanently irrigated*, *Agro-forestry*, *Olive groves* and *Sclerophyllous vegetation* exhibit at least two different types tested within the framework (forest, crop, grassland) or substantial differences if mixed with another land-cover type (irrigated vs non-irrigated). On the contrary, the CORINE class *Broad-leaved forest* should be subject to relatively little land-use change and a strong seasonal vegetation signal (amplitude) which showed to diminish differences between methods. This is largely reflected in Figure 3.10 where the former show large inter-quantile ranges within and generally large differences between methods. The latter is subject the lowest inter-quantile ranges and also exhibits a consistent trend towards earlier green-up.

3.4.2 Satellite data trend estimates

Differences between extracted d_{GU} and observed d_{BB} were within the anticipated ranges (1-25 days). A few studies analysed the relationship between extracted d_{GU} and observed d_{BB} . Liang et al. (2011) for example related 'Full Budburst' – which is equivalent to d_{BB} of the DWD – to d_{GU} using the method of Zhang et al. (2003) (method L_g in this paper). They found differences of ~ 1 day when using the EVI (Enhanced Vegetation Index) and up to 2 weeks (earlier) for some years when using the NDVI. In this case d_{GU} was determined where change rates in the curvature of fitted sigmoidal functions exceeded local minima or maxima. This corresponds to substantially earlier d_{GU} compared to the thresholds applied in this paper, i.e. the initial NDVI increase in spring. Thayn (2012) related vegetation cover fraction (VCF) of sweet corn to respective d_{GU} – again using the methods of Zhang et al. (2003). They found at VCF $\sim 10\%$ or 2 cm height the majority of extracted d_{GU} . Ryu et al. (2012) detected 5 days difference between d_{GU} and d_{BB} for an oak-savanna ecosystem. Soudani et al. (2008) found a difference of 8.5 days between d_{GU} and d_{BB} for deciduous tree species in France when 90% of all trees have buds open of at least 20% of their crown. They also used a sigmoidal function but with d_{GU} at the inflexion point. Thresholds for d_{GU} extraction used in this paper all refer to later developmental stages than the studies mentioned. Consequently, one would expect a difference between d_{GU} and d_{BB} of at least 2-3 weeks which are approximately the extracted differences between ground observations and logistic/sigmoidal functions.

Trend estimates are in line with previous studies using one methodology and averaging all land-cover types. Julien and Sobrino (2009) found a global d_{GU} advancement of 0.38 days/year and for Europe ~ 3 days/year using AVHRR data (1981-2003). Zhang et al. (2013) used MODIS, AVHRR and SPOT observations and concluded a spring advancement of ~ 1

days/year for the Tibetan Plateau. Stöckli and Vidale (2004) determined a spring advancement for Europe from 1982-2001 of 0.54 days/year based on AVHRR observations. Delbart et al. (2006) extracted an earlier d_{GU} in boreal regions using AVHRR and SPOT between 1982-1991 (1 days/year) but a delay of 0.3 days/year between 1993-2004. Hamunyela et al. (2013) analysed the length of the vegetation period in Europe with MODIS for 2001-2011 and reported an increase of 0.1-1.5 days/year primarily caused by a later senescence (green-up unchanged). d_{GU} trends even for a single land-cover class (deciduous forest) were ambiguous with 40% positive and 60% negative trend. Jeong et al. (2011) found 5.2 days spring advancement between 1982-1999 but a stable seasonal timing between 2000-2008. These studies used a single method (e.g. logistic functions, Fourier transformation, polynomials) for NDVI temporal profile approximation.

3.4.3 Comparison to studies using method ensembles

Results of studies using method ensembles show a different picture. White et al. (2009) found no evidence for an earlier spring arrival in North America. Fu et al. (2014) detected a phenology shift towards earlier green-up until the year 2000, which was in line with ground observational records. However, satellite-based trends differed from 2000 onwards, exhibiting trends towards later spring onset. Wang et al. (2015) found a linear change towards earlier d_{GU} in the Northern Hemisphere but still considerable trend differences in space for North America but also Eurasia. Similar results were extracted from (Cong et al., 2013) showing substantial trend differences between methods and also across land-cover classes for a single method. These findings strongly indicate a different suitability of methods employed which was also shown by White et al. (2014) even when using logistic functions over deciduous forest only.

As shown in the theoretical part of our study a method's performance to extract trends is affected by the shape of the NDVI' seasonal evolution. Land-cover types like grassland exhibit a more moderate NDVI increase in spring compared to crops or deciduous forest, which results in potentially blurred trend estimates for most approximating methods. A mix of green-up signals showed to have similar effects. This limits the application for some algorithms presented here and used in former works and will translate into different trend magnitudes. We demonstrated this issue also by introducing a goodness-of-fit measure across applied methods. Differences between methods (approximating vs interpolating) and land-cover classes

were evident. Lower fit quality will generally increase the spread of subsequently computed trends (all pixels of one land-cover) and potentially also modify the respective mean trend.

Despite those varying suitability method ensembles used here indicated a moderate spring advancement also across land-cover types in contrast to other studies. This could originate from applying an identical filtering procedure (BISE) prior to fitting interpolating or approximating functions to the remaining NDVI values. Methods of other ensembles applied other filtering approaches which might amplify trend differences stemming from disparate suitability mentioned above.

3.4.4 The satellite's spatio-temporal resolution

The spatio-temporal resolution of utilised satellite data had certainly an effect on extracted phenological metrics compared to previous studies. Here, we used daily NDVI observations. Lower temporal resolutions might mask some phenological transitions (Fisher et al., 2006). One has to assume that data of 1 km resolution is less disturbed by potentially confounding factors such as mixed pixels than the usually applied 8 km resolution of the GIMMS dataset. Consequently, the effects shown in the theoretical part of our study should be pronounced more in studies relying on, e.g. the GIMMS dataset. The likelihood to capture temporally changing surface elements increases with a coarser spatial resolution. Trend estimates should consequently be less affected by changes in land-use practice (Reed, 2006) or underlying land-cover change when using higher resolution data. Even small fractions of these elements can alter the signal received by the satellite (Doktor et al., 2009) and can affect trend estimates.

In some regions of intensive agriculture, namely the upper Rhine valley (Germany/France) and the Po valley (Italy) relative areas of cultivated crops changed substantially and sustained over the past decades with e.g. maize replacing wheat as the dominating crop. As a result, the landscape exhibits more the seasonal evolution of a summer crop which shifted the satellite's received green-up signal. Such clear and large-scale changes are picked-up by NOAA AVHRR as identified positive trends in those regions. At smaller scales, however, such changes should blur trend estimates and increase respective uncertainties as the mixture of underlying land-cover / land-use is unstable for a larger pixel number at 8 km than at 1 km spatial resolution. Here, just one example for crop rotation influence is given but other land-use changes with respect to impervious surfaces, deforestation or afforestation impact

the surface signal similarly.

3.4.5 Ground truth & modelling

The DWD ground observations used here are well suited to evaluate phenological metrics from remotely sensed imagery. First they constitute a dense network over a fairly large geographical area providing time-series of ~ 20 or more years. Furthermore, observations include wild-growing species (forest) which is important as only stationary ground observations can reasonably be synchronised with satellite observations. The combined time series including outlier detection assure robust trend estimates also in case of highly variable phenological data. Combined time series is also an advantage when training the applied phenological model (PIM). Similar trend estimates for all three approaches (satellite data, ground observations and model) suggest robustness of the presented framework in this study.

3.4.6 Sensor inter-calibration and potential false trends

The applied MEDOKADS dataset covers the operational time span of several satellites (NOAA 11, NOAA 14, NOAA 16, NOAA 17). Hence, sensor inter-calibration is an important issue and has affected AVHRR time series' robustness until the generation of the GIMMS 3G dataset. For example Shen et al. (2011) and Zhang et al. (2013) discovered diverging trends between MODIS and GIMMS based analysis. MEDOKADS inter-sensor calibration has been performed independent of past GIMMS datasets. Nevertheless, we checked the plausibility of extracted phenological metrics for the years 2001-2007 against i) trends based on Terra MODIS (MOD09GQ) and ii) phenological ground observations.

Trend estimates for areas of deciduous forest in Germany (2001-2007) are similar comparing Terra MODIS and AVHRR with ~ 0.4 and ~ 0.6 days/year respectively towards earlier green-up when using local thresholds. AVHRR yearly mean green-up dates of broadleaf forest in Germany are also in line with respective ground observations of the DWD: trend ~ 0.8 days/year and correlation > 0.7 . It is therefore unlikely that the MEDOKADS dataset does suffer from inter-sensor calibration problems as past GIMMS datasets. Thorough statistical analysis as outlined in de Beurs and Henebry (2004) are nevertheless necessary to substantiate this further but was out of scope for this study.

Extracted trend magnitudes, however, can for example be explained by i) trends (positive or negative) in NDVI maximum/minimum values and ii) high between year variation of NDVI maximum/minimum values. These issues will be discussed in the following as they can cause threshold dependent differences in detected trend magnitudes. False green-up trends might be detected based on artificially higher or lower NDVI maximum values. Changing aerosol contents over the analysed time period could for example lead to an NDVI increase (or decrease) which then affects extracted phenological metrics. Central and Eastern Europe recorded decreasing aerosol emissions after 1990 due to the collapse of many polluting industries. The eruption of Mount Pinatubo on the other hand lead to globally generally increasing aerosol contents (Bluth et al., 1992). Another factor is the effect of CO₂ fertilisation which has been found to translate into a greening of the vegetation (Donohue et al., 2013) and therefore higher NDVI introducing an artificial d_{GU} delay.

In order to assess the affect of changes in the NDVI range we computed the trends of extracted maximum NDVI values. We found a mean trend towards higher NDVI maximum values of ~ 0.006 days/year for broadleaf forest in Germany that was not statistically significant. Nevertheless, an average increase of NDVI maximum values of ~ 0.11 over the analysed period can be expected for some part of the study area. According to Duchemin et al. (1999) the average NDVI increase for NOAA AVHRR in temperate forests in spring is ~ 0.013 days/year. This translates into a d_{GU} shift of ~ 8 days and will consequently affect trend estimates thereof. Eastman et al. (2013) also observed a changing seasonality analysing global AVHRR data from 1982-2011: an uniform NDVI increase for broadleaf forests, a higher NDVI amplitude in grasslands and a NDVI maximum increase in Tundra regions. A higher NDVI amplitude in grassland could hint towards intensification and would artificially delay extracted d_{GU} dates as found in this study.

DOY determination of NDVI minimum values is less reliable because NDVI differences in winter time are relatively small over dormant vegetation and furthermore error prone due to large BRDF effects and wet soil influences. Consequently, a small NDVI difference can shift the respective DOY without actual physical differences of the observed vegetation. Therefore, trends (and variation) of DOY minimum values were not further investigated. Nevertheless, a potential source explaining differences between methods stems from difficulties to reconstruct the true NDVI profile in winter months: Snow cover or wet vegetation can cause high NDVI values even when almost no green/active vegetation is present (Zhang et al., 2006). Interpolating methods such as *Li* can compensate the selection of potentially false NDVI values in winter because the method will match lower subsequent NDVI values

before the start-of-season. Approximating methods such as *Lg* do, in contrast, not exhibit values lower than the initial value before the maximum of that function is reached. The computation of that initial value also differs between the approximating methods. Consequently, high – potentially wrong – NDVI values in winter could account for differences between interpolating and approximating methods.

The choice of the time period (available time series) can exert a crucial influence on trend magnitudes. For testing whether other start/end points change the trend magnitude we computed trend estimates for the time periods of 1990-2007 and 1989-2006 using the method *Li*. Only negligible differences compared to the original time period, i.e. 1989-2007, were found. We therefore conclude that i) the time series in use is not particularly affected by extreme years (at the beginning/end) and ii) trend estimates are robust and representative.

3.4.7 Recommendations

This study provides a theoretical framework that explains the sometimes ambiguous results from past studies on derived green-up trends using satellite times series of relatively coarse spatial resolution. Results stemming from the application of method ensembles on simulated satellite times series suggest that the derivation of phenological metrics – and respective trends – is influenced by the underlying land-cover (even if homogeneous) and also the number and relative areas of different land-cover types mixed in a satellite pixel. These results are confirmed by 1) the goodness-of-fit statistics of the method ensemble on real satellite data and also 2) by more robust trend estimates of land-cover types which exhibit a strong seasonal vegetation signal and are just moderately affected by land-use / land-cover change: broadleaf forests.

Methods to derive phenological metrics should be selected depending on the analysed land-cover type and the land surface's spatial heterogeneity. Hence, simultaneously employing detailed land-use classification schemes. Our results favour the use of interpolation approaches to reconstruct the seasonal NDVI profile as opposed to approximating methods – with the exception of pixels where two land-cover types are present with similar area ratios. (Double) logistic functions outperform other approximating methods at the level of interpolating algorithms.

Despite the above mentioned differences within the method ensemble we could detect a moderate spring advancement for pseudo-invariant land-cover classes such as deciduous forest

across all methods. This might also be due to using identical filtering procedures on raw NDVI observations and employing data with a spatial resolution of 1 km rather than 8 km. If possible large scale trend estimates should be evaluated against estimates stemming from an pseudo-invariant land-cover (homogeneous pixels, stable geolocation) from within that area. When comparing satellite imagery based phenology to respective ground observations a pre-processing of the latter is necessary, i.e. using combined time series.

Data accessibility

Phenological ground observations

The phenological ground observations used here to built the combined time series can be accessed via the German Weather Service (ftp-cdc.dwd.de/pub/CDC/observations_germany/phenology/annual_reporters/wild/).

Satellite imagery

- 1) The NOAA AVHRR MEDOKADS dataset can be requested via www.ufz.de/drpf, the data investigation portal of the Helmholtz Centre for Environmental Research - UFZ
- 2) MODIS Terra data can be downloaded via e4ftl01.cr.usgs.gov

Meteorological data

Daily mean temperatures (for driving the phenological model PIM) were obtained from Web-Werdis, the Web-based weather request and distribution system of Germany's national meteorological service (Deutscher Wetterdienst, DWD). The service was discontinued in 2018, data can now be obtained from the Climate Data Center opendata.dwd.de.

Land-cover data

CORINE can be downloaded freely from the European Environmental Agency (EEA, European Environment Agency (2013)).

4

Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors

Authors Maximilian Lange¹, Benjamin Dechant^{1,2}, Corinna Rebmann³, Michael Vohland³, Matthias Cuntz^{2,4} and Daniel Doktor¹

Affiliations ¹ Department Computational Landscape Ecology, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany

² Department Computational Hydrosystems, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany

³ Leipzig University, Institute for Geography, Geoinformatics and Remote Sensing, Johannisallee 19a, 04103 Leipzig, Germany

⁴ INRA, Université de Lorraine, UMR1137 Ecologie et Ecophysiologie Forestière, 54280 Champenoux, France

Published in: Sensors 17 (8), 1855
DOI:10.3390/s17081855

Publishing date: 11 August 2017

Acknowledgements The authors wish to thank Sebastian Gimper and Hendrik Zöphel from the Helmholtz-Centre for Environmental Research-UFZ, Department Computational Hydrosystems, for their technical support regarding sensor installation, as well as Hannes Hertel for his support regarding sensor programming. We would like to thank the Helmholtz Association and the Federal Ministry of Education and Research (BMBF) for providing infrastructure in the framework of TERENO (Terrestrial Environmental Observatories). This project was supported by the Federal Ministry for Economic Affairs and Energy Germany (Grant Number 50EE1218) and has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 641762. The MODIS MOD09GQ and MYD09GQ data products were retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool.

Author Contributions All authors contributed significantly to the work presented in this paper. Daniel Doktor and Maximilian Lange designed the present study and wrote the manuscript. Maximilian Lange installed the sensor systems together with Benjamin Dechant, implemented the algorithms used in this study, analyzed and processed the data. Benjamin Dechant installed the sensors together with Maximilian Lange and calibrated the instruments. Corinna Rebmann and Matthias Cuntz provided the infrastructure. Corinna Rebmann organized the installation of the sensor systems. Michael Vohland helped with designing the manuscript and writing. Matthias Cuntz and Daniel Doktor initiated the extension of 'Hohes Holz' as a satellite validation site. All authors edited the manuscript.

Abstract Quantifying the accuracy of remote sensing products is a timely endeavor given the rapid increase in Earth observation missions. A validation site for Sentinel-2 products was hence established in central Germany. Automatic multispectral and hyperspectral sensor systems were installed in parallel with an existing eddy covariance flux tower, providing spectral information of the vegetation present at high temporal resolution. Normalized Difference Vegetation Index (NDVI) values from ground-based hyperspectral and multispectral sensors were compared with NDVI products derived from Sentinel-2A and Moderate-resolution Imaging Spectroradiometer (MODIS). The influence of different spatial and temporal resolutions was assessed. High correlations and similar phenological patterns between in situ and satellite-based NDVI time series demonstrated the reliability of satellite-based phenological

metrics. Sentinel-2-derived metrics showed better agreement with in situ measurements than MODIS-derived metrics. Dynamic filtering with the best index slope extraction algorithm was nevertheless beneficial for Sentinel-2 NDVI time series despite the availability of quality information from the atmospheric correction procedure.

4.1 Introduction

Quantifying the accuracy of satellite-based remote sensing products is an important task that is increasingly timely given the rising number of Earth observation missions. Different measurement strategies are pursued at established validation sites, including continuous in situ (spectral) measurements, airborne remote sensing campaigns at the landscape level, as well as the collection of independent satellite sensor data (Justice et al., 2000; Morisette et al., 2006). Linking small-scale observations with large-scale satellite products is crucial to understand the relationships between optical information and ecological, as well as physiological processes (Gamon, 2015). This thus facilitates prediction and mapping of processes at regional and global scales, such as terrestrial carbon assimilation (Balzarolo et al., 2016). Automatic continuous spectral measurements are bridging the gap between continuous micro-meteorological measurements and remote-sensing products by providing optical information about vegetation processes with high temporal resolution (Balzarolo et al., 2011; Porcar-Castell et al., 2015). They are thus helpful to improve the understanding of satellite measurements.

Phenology plays a central role in the estimation of vegetation-dependent processes (Lange et al., 2016). Research on vegetation phenology uses phenological and spectral ground observation networks, phenological modeling, eddy covariance towers and satellite-derived imagery to assess and monitor vegetation status and dynamics (White et al., 2009; Vanbeverem et al., 2016). Several studies are dealing with the question of which vegetation index to choose for satellite phenology depiction (Jin and Eklundh, 2014). Although the Normalized Difference Vegetation Index (NDVI) shows saturation effects in dense forest canopies (Jöns-son et al., 2010), it is one of the most widely-used indices for this purpose, mainly due to data availability and its robustness against noise and varying illumination geometries (Doktor, 2008; Jin and Eklundh, 2014). A large number of studies (Heumann et al., 2007; Karlsen et al., 2008; Myneni et al., 1997; Zhou et al., 2001) mapped land surface phenology from satellite NDVI data. These studies outlined two main problems in the derivation of phenology

from satellite data: firstly, the low spatial resolution of satellite data, e.g., from the NOAA Advanced Very High Resolution Radiometer (AVHRR) and the NASA Moderate-resolution Imaging Spectroradiometer (MODIS). This results in mixed pixels, integrating phenological signals of different vegetation and land cover types, e.g., deciduous forests and nearby grasslands. New satellite generations, such as the Copernicus Sentinel-2, may overcome this issue with higher spatial resolution. Secondly, non-vegetational effects may alter the received phenological signal. These effects include atmospheric conditions, snow cover (Delbart et al., 2005; Jönsson et al., 2010), soil-wetness (White et al., 2009), viewing geometry and illumination conditions (Stöckli and Vidale, 2004; Eklundh et al., 2011), as well as the distorted signal under overcast conditions. Noise introduced by these effects can be eliminated by using methods like Maximum Value Composite (MVC) (Holben, 1986) or using dynamic filtering, such as the Best Index Slope Extraction (BISE) (Viovy et al., 1992). While MVC is effective at reducing cloud and viewing condition effects, it can include outliers with higher NDVI. Short-term vegetation-state changes might be masked using a long composition period, while a short period retains noise (Viovy et al., 1992). BISE excludes outliers, and its parametrization options allow for the removal of sudden decreases in NDVI, mainly due to non-vegetational effects. This, however, also excludes sudden decreases in NDVI due to vegetational processes, followed by rapid regrowth (Viovy et al., 1992). Several approaches are used to fill gaps and smooth the phenological time series after the removal of noisy measurements: interpolation methods, signal filter or model fitting algorithms (Doktor, 2008). Subsequently, phenological metrics are extracted using different approaches: global or local thresholds, conceptual-mathematical models (White et al., 2009) or local extrema of transition rates (Zhang et al., 2003). Here, we used BISE to filter NDVI data, followed by a simple linear interpolation to remove gaps in the time series, and finally, we extracted phenological metrics with a local threshold method.

The aim of this study is to validate different satellite phenology products by comparing them to multi-sensor ground measurements. This also requires a detailed description of the site's instrumentation, as well as information on calibration, but the main focus of this study is not on the latter. Details on calibration will be described in a separate publication (Dechant et al., 2017). In the following, we briefly introduce measurement approaches commonly applied at spectral ground observation sites.

Compact multispectral sensors such as the SKYE SKR 1860 series (Skye Instruments Ltd., Llandrindod Wells, Powys, U.K.) are comparably low-cost and fairly straightforward to set up and maintain (Porcar-Castell et al., 2015). Typically, no optical fibers are required, and

data can be logged with data loggers commonly used at eddy covariance sites. Therefore, this type of instrument is widely used (Pacheco-Labrador and Martín, 2015). Hyperspectral systems, in contrast, are usually more expensive and require computers to control the timing of measurements and instrument settings (Porcar-Castell et al., 2015). Furthermore, it is common that fiber optics are used to connect the spectrometer to the point of measurement (Drolet et al., 2014; MacArthur et al., 2014; Sakowska et al., 2015), and this introduces additional potential for calibration issues on top of spectrometer calibration itself (Porcar-Castell et al., 2015). However, they provide more detailed spectral information. Both sensor types are potentially affected by issues of long-term continuous field measurements (Balzarolo et al., 2011; Porcar-Castell et al., 2015), such as degradation and sensitivity to temperature and humidity. While other studies focused mainly on either hyperspectral (Meroni et al., 2011; Drolet et al., 2014; MacArthur et al., 2014; Pacheco-Labrador and Martín, 2015) or multispectral sensors (Eklundh et al., 2011; Balzarolo et al., 2016), we used both types of sensors in order to assess measurement quality issues introduced by environmental conditions and instrument settings (Pacheco-Labrador and Martín, 2015), which might be relevant for the validation of satellite-derived phenological profiles.

In order to compute a reflectance factor, later used for the calculation of NDVI, it is necessary to measure the upwelling and downwelling radiation fluxes. This can be achieved with different setups (Porcar-Castell et al., 2015): systems using one sensor have to measure a reference in sequence to each measurement of the target surface, e.g., a white reference panel (Sakowska et al., 2015) or downwelling irradiance through the rotation of the fore optics of the sensor (Meroni et al., 2011). This approach introduces considerable time delays between downwelling irradiance and upwelling radiance measurements and therefore increases measurement uncertainties, particularly under unstable illumination conditions (Sakowska et al., 2015). The moving parts are an additional potential error source in a long-term outdoor setup. Furthermore, the approach is restricted to a single fore optic for both target and reference measurements, which is suitable only for measurements of bi-hemispherical reflectance factors in practice.

Measuring downwelling irradiance and upwelling radiance quasi-simultaneously with a dual-field-of-view (DFOV) setup and different FOV fore optics (Eklundh et al., 2011; Drolet et al., 2014; Pacheco-Labrador and Martín, 2015) removes the restriction of observing bi-hemispherical reflectance factors. Furthermore, simultaneous acquisition for both FOVs is possible in the case of the use of two sensors, while the time delay depends on sensor characteristics and instrument settings such as integration time if only one sensor is used

(MacArthur et al., 2014). However, challenges due to spectral shifts caused by bifurcated fibers (one sensor) and cross-calibration of wavelength scale (two sensors) may arise for DFOV hyperspectral measurements. We used the DFOV approach with two sensors for multispectral measurements and with a single sensor and a bifurcated fiber for hyperspectral measurements. Hyperspectral- and multispectral-based NDVI values (spatial extent on the order of 10 m) were compared with NDVI products derived from MODIS Aqua and Terra (250-m ground resolution), as well as Sentinel-2 (10- and 20-m ground resolution) over the course of two vegetation periods. We evaluated whether dynamic filtering procedures, commonly used with NOAA AVHRR and MODIS, are now obsolete with new generation satellite data. Finally, the influence of varying spatial resolutions of the different datasets on extracted phenological metrics was also assessed.

4.2 Materials and Methods

4.2.1 Study Site and Sensor Setup

We collected data at the forest site ‘Hohes Holz’ (52.08° N 11.22° E), which is situated in central Germany near the Harz Mountains in a temperate climate (Wollschläger et al., 2016). The site is part of the interdisciplinary and long-term research programs Tereno (Terrestrial Environmental Observatories (Bogena et al., 2011)) and ICOS (Integrated Carbon Observation System (ICOS ERIC, 2017)). The ecosystem of ‘Hohes Holz’ is a deciduous forest with a size of around 15 km². The vegetation within one MODIS pixel around the tower consisted of 50% oak, 45% beech, 2.5% birch and 2.5% clearings, whereas the area within 30 m around the tower had a higher amount of clearings (10%) and birch (4%) and less oak (44%) and beech (42%). The eddy flux tower has a height of 50 m, which is about 20 m above the canopy. We installed one hyperspectral sensor and two multispectral sensors at the site: a QE65000 (Ocean Optics, Dunedin, FL, USA) and an SKR1850 4-channel sensors (Skye Instruments Ltd., Llandrindod Wells, Powys, U.K.).

Data logging of the multispectral sensors is performed with a Campbell Scientific CR1000 data logger with a sampling frequency of 0.1 Hz. The upward facing multispectral sensor is set up with a hemispherical cosine diffuser. The downward facing sensor is used for narrow-angle measurement (25° FOV) of vegetation with a south-orientated (view azimuth of 180°) off-nadir view angle of 22.5°. Central wavelengths and the full-width at half-maximum

(FWHM) were chosen according to Sentinel-2 Channels 4–6 and 8a (see Table 4.1).

Band	CW (nm)	FWHM (nm)
1	654	37.50
2	708	8.75
3	739	9.50
4	858	9.50

Table 4.1: Configuration of the multispectral sensor. Central wavelengths (CW) and the full-width at half-maximum (FWHM) bandwidth are given for each band.

The hyperspectral sensor uses scientific-grade back-thinned detectors (S7031-1006, Hamamatsu Photonics, Hamamatsu City, Shizuoka Pref., Japan) and was set up with a spectrometer slit width of 10 μm , allowing an FWHM of 1.58 nm and a spectrometer grating allowing for observations in the range of 398 nm–1174 nm with 1030 channels (QE65000, Ocean Optics, Dunedin, FL, USA). A 400- μm bifurcated fiber optic cable (QBIF400-VI, Ocean Optics, Dunedin, FL, USA) is used as the dual input channel, where each of the two cable ends is connected with a shutter (INLINE-TTL, Ocean Optics, Dunedin, FL, USA). Each shutter is connected with a 400- μm fiber optic cable (QP400-15-V, Ocean Optics, Dunedin, FL, USA). A cosine corrector (CC-3-UV-S, Ocean Optics, Dunedin, FL, USA) is attached to the upward facing cable end for irradiance measurements. The downward facing cable end is used without fore optics with a field of view of 25° to measure upwelling radiances. The optics are mounted 20 m above the canopy with a south-orientated (view azimuth of 180°) off-nadir view angle of 22.5° (see Figure 4.1b) on a boom of the eddy flux tower. This results in an observed canopy area of around 72 m^2 .

The sensor and shutters are controlled via LabView (National Instruments, Austin, TX, USA). Shutters open and close alternately to select the signal from upward- and downward-facing fibers to measure downwelling irradiance and upwelling radiance sequentially. Integration time was set to two seconds between April and September and four seconds between October and March. Dark current is determined with closed shutters before and after each measurement. This procedure is repeated three times every ten minutes and takes 24 seconds (three times eight seconds) between April and September and 48 seconds (three times 16 seconds) between October and March. The whole setup (see Figure 4.2) is situated in a weather-proof box (see Figure 4.1a) with a temperature-driven ventilation system. USB devices are connected to a USB-to-Ethernet server to enable remote control. The data logger is connected with Ethernet, and data are transferred automatically once a day onto a network

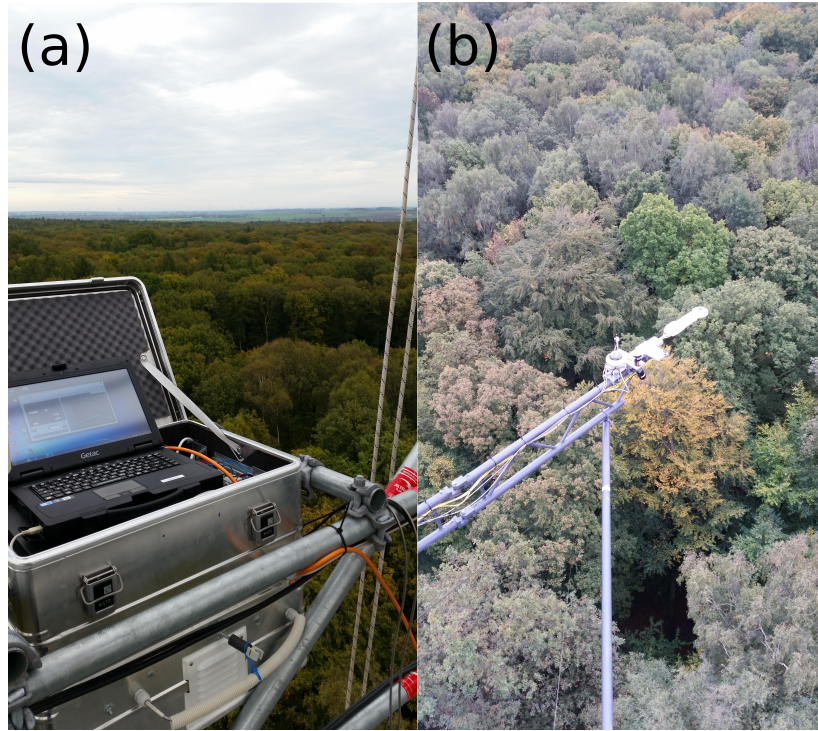


Figure 4.1: The weather-proof box is mounted at the outside of the eddy flux tower (a). The multi-spectral sensor and optics of the hyperspectral sensor are mounted on a boom of the tower at 49-m height (b).

storage system.

4.2.2 Generation of NDVI Products from Ground-Based Spectral Measurements

Hyperspectral solar irradiance data were cross-calibrated once by using measurements of solar irradiance with an ASD FieldSpec 4 spectrometer (Analytical Spectral Devices Inc., Boulder, CO, USA) and a Spectralon panel. Calibration performance was assessed by re-sampling the QE65000 spectrum to the lower spectral resolution of ASD FieldSpec. High accuracy and precision were achieved when comparing the corresponding two spectra of solar irradiance ($r^2 = 0.99$, relative $RMSE = 0.03$ in the spectral range 440–1000 nm). Hyperspectral upwelling radiance observations were calibrated by using the downwelling irradiance calibration and a Spectralon panel. The stability of wavelength positions was irregularly

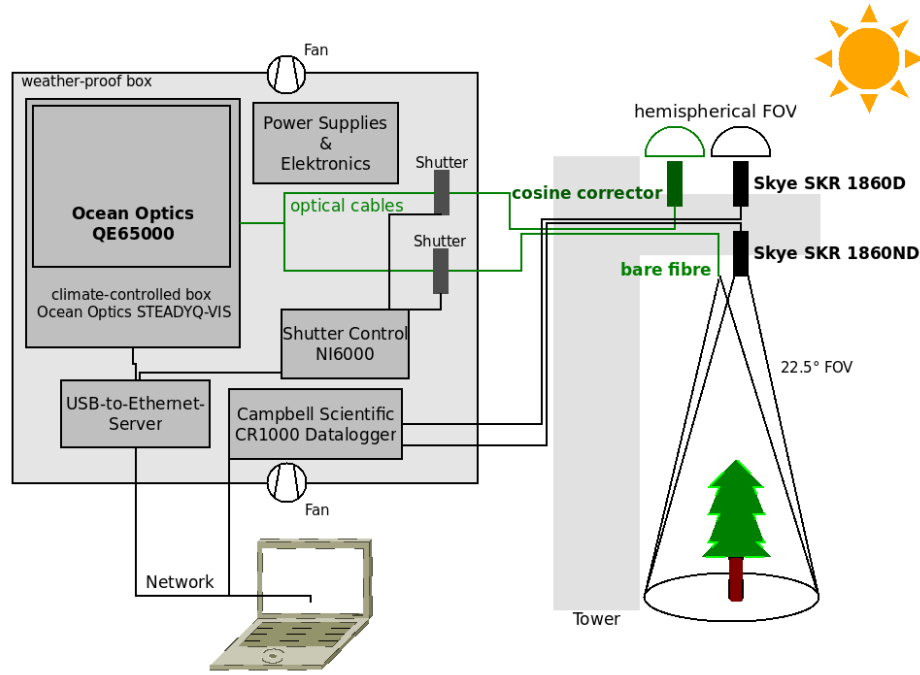


Figure 4.2: Schematic overview of sensor setup for unattended hyperspectral and multispectral measurements. Electronics and the hyperspectral sensor are installed in a weather-proof box and connected to the network. Fore optics and multispectral sensors are mounted on a boom of a tower above the forest.

checked with a mercury argon calibration light source (HG-1, Ocean Optics, Dunedin, FL, USA). Multispectral data were radiometrically corrected by using the channel sensitivities given by the manufacturer. Manufacturer calibration was repeated every two years (2014 and 2016). Long-term stability was assessed by comparing against shortwave pyranometer measurements (CNR4 net radiometer, Kipp&Zonen B.V., Delft, The Netherlands) from the eddy flux tower. We verified CNR4 calibration stability by comparing against other sensors of the same type at sites located not far from the forest site. Channel 1 of the upward-looking multispectral sensor did not give correct observations. Its values were hence replaced with appropriately-scaled values of Channel 2. The validity of this approach was assessed both with hyperspectral observations at the same site and comparable multispectral observations at a different site (Dechant et al., 2017).

Hyperspectral data were normalized by integration time, and the electronic dark current was subtracted before correction. Hyperspectral observations were transformed into multispectral signals by multiplying them with the corresponding spectral response curves of

the multispectral sensor and integrating over the wavelengths, for both downwelling irradiance and upwelling radiance before calculating reflectance factors. Measurements are taken three times (hyperspectral sensor) every ten minutes and 60 times within ten minutes (multispectral sensors); we refer to this further as the measurement-cluster. For each cluster, the absolute sum of irradiance change per channel ($Irr_{i,change}$) and the mean irradiance per channel (\overline{Irr}_i) are calculated. Measurement clusters are rejected if $Irr_{i,change}$ exceeds $0.02 \times \overline{Irr}_i$ (hyperspectral data) or \overline{Irr}_i (multispectral data). Hence, measurements are selected that have been acquired under steady conditions, i.e., mainly clear sky conditions. Subsequently, only measurements between 11:00 a.m. and 01:00 p.m. UTC are considered. Non-vegetational effects usually decrease NDVI values, but highly diffuse illumination conditions illuminating shaded areas and therefore producing directional effects, for example, might also increase NDVI values. Such effects are rather difficult to eliminate. Here, we used sunshine pyranometer data (SPN1, Delta-T Devices Ltd., Cambridge, U.K.) to exclude measurements with a direct to diffuse solar radiation ratio (further referred to as $DDSR$) lower than the empirically determined value of 3, to exclude the effects of varying illumination conditions that can affect NDVI (Damm et al., 2015). Only upwelling radiance measurements made simultaneously to selected downwelling irradiances were used in order to compute reflectance factors: $Refl_i$, $i = 1, \dots, 4$ (see Equation (4.1)), where π stands for corrections of the hemispheric angle (Nicodemus et al., 1977).

$$Refl_i = \pi \times \frac{Rad_i}{Irr_i} \quad (4.1)$$

Mean reflectance factors of each measurement cluster serve the generation of NDVI data with $RED = Refl_1$ and $NIR = Refl_4$ according to Equation (4.2). The resulting datasets are further referred to as $NDVI_{multi}$ and $NDVI_{hyper}$ and contain 1525 and 1447 data points for ‘Hohes Holz’, respectively.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (4.2)$$

4.2.3 Satellite Data and Respective NDVI Products

We used spectral data from MODIS datasets MYD09GQ (Vermote, E. and Wolfe, R., 2015b) and MOD09GQ (Vermote, E. and Wolfe, R., 2015a) with a spatial resolution of 250 m, as

well as Sentinel-2A Level 1C (L1C) data (Copernicus Sentinel data, 2017) with spatial resolutions of 10 and 20 m. Sentinel-2A L1C data were atmospherically-, terrain- and cirrus-corrected with Sen2Cor (Version 2.3.0) (Mueller-Wilm et al., 2016) and ATCOR 2/3 (Atmospheric & Topographic Correction for Small FOV Satellite Images, RESE®, Version 9.1) (Richter and Schl pfer, 2016) in order to generate Level 2A (L2A) data. Both processing tools generate scene classifications to distinguish between clear pixels and clouds, shadows and saturated pixels. We continued the study with ATCOR 2/3 (and without Sen2Cor) due to the higher number of successfully-processed scenes with clear sky conditions. ATCOR 2/3 could process 47 out of 78 scenes without error messages. The remaining scenes could not be processed due to the solar zenith during acquisition exceeding 70°, software problems or less than one percent of clear pixels. ATCOR 2/3 detected 22 scenes as ‘clear’ around the site ‘Hohes Holz’, while eleven of them showed clouds or shadows after manual examination.

NDVI (see Equation (4.2)) is calculated from Bands 1 (RED) and 2 (NIR) of products MYD09GQ and MOD09GQ, resulting in the datasets $NDVI_{aqua}$ and $NDVI_{terra}$. Sentinel-2A NDVI is calculated from Bands 4 (RED) and 8 (NIR, 10-m resolution) or 8a (NIR, 20-m resolution) of Sentinel-2A L2A data. These datasets are further referred to as $NDVI_{s2a,10}^{atcor}$ and $NDVI_{s2a,20}^{atcor}$, respectively (Figure 4.3).

4.2.4 NDVI Post-Processing and Phenological Metrics Extraction

Methods for NDVI post-processing and extraction of phenological metrics were implemented in an R-package (R Core Team, 2014) for phenological data analysis called ‘phenex’ (Lange and Doktor, 2017). This includes the Best Index Slope Extraction (BISE) (Viovy et al., 1992) for preprocessing and methods to model phenological time series.

BISE is a dynamic filter algorithm for time series. Starting at the first date of the time series (i.e., DOY = 1 for each year), BISE accepts the next data point if it has a higher value than the previous one. Lower values are accepted if there is no point in a pre-defined time period (‘sliding period’) with a higher value. Higher values are rejected if they exceed a pre-defined threshold $thres_{inc}$, i.e., the allowed increase of NDVI per day. NDVI fluctuations from natural surfaces greater than 0.1 per day are usually attributed to data errors.

We applied BISE on $NDVI_{aqua}$ and $NDVI_{terra}$ with a sliding period of 30 days and an allowed NDVI increase of 10% per day ($thres_{inc} = 0.1$) in order to remove outliers and eliminate non-

Sentinel-2A NDVI time series (Hohes Holz 2016)

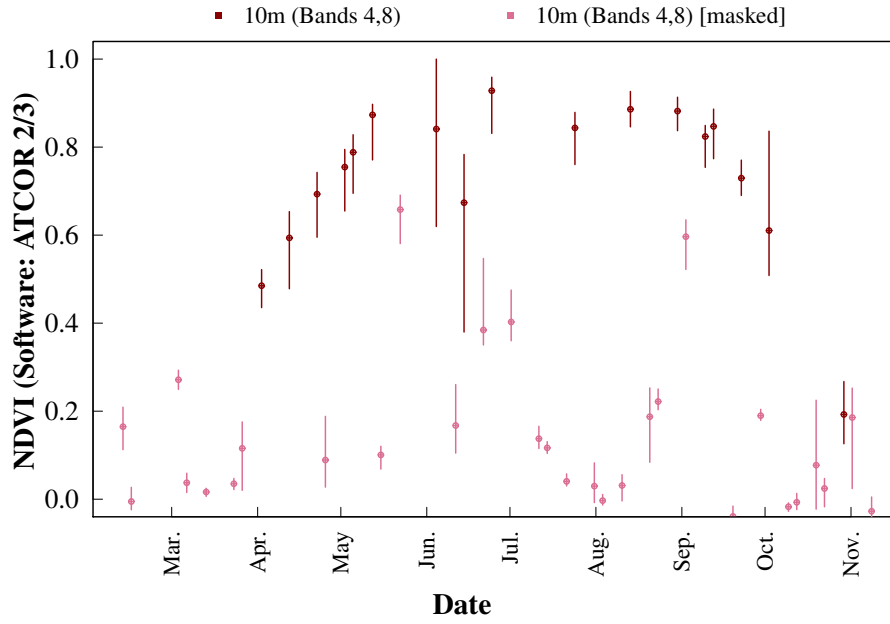


Figure 4.3: Sentinel-2A NDVI time series from 2016 with 10-m resolution processed with ATCOR 2/3. The lines show the distribution of NDVI values of pixels in an area within 30 m around the eddy flux tower. The dot marks the respective median value. The NDVI of clear pixels are shown in dark colors, whereas pixels with clouds, shadows or saturation are shown in pale colors (masked values).

vegetational effects, e.g., cloud cover. The resulting datasets were stored as $NDVI_{aqua}^{bise}$ and $NDVI_{terra}^{bise}$.

The next step is the reconstruction of daily NDVI values based on BISE-selected NDVI observations ($NDVI_{multi}$, $NDVI_{hyper}$, $NDVI_{aqua}^{bise}$, $NDVI_{terra}^{bise}$ and $NDVI_{s2a,10}^{atcor}$). Here, we used linear interpolation; however, other approaches such as fitting a Gaussian function are also commonly used (White et al., 2009). Phenological metrics extraction is subsequently conducted on the reconstructed NDVI profile and includes green-up and senescence dates, as well as maximum and minimum NDVI and their date of occurrences. Green-up and senescence dates are defined as dates when NDVI values reach a threshold of 55% between the minimum and maximum NDVI values (Doktor and Lange, 2017; Lange and Doktor, 2017). The threshold was determined empirically by comparing satellite-derived green-up dates from deciduous broadleaf forests all over Germany to respective ground observations of the German Weather Service (Deutscher Wetterdienst, DWD) acquired from the Plant-Phenological

Online Database (PPODB, www.phenology.de (Dierenbach et al., 2013)). This database provides around 1500 observations per species, phenological phase and year. The standard deviation of phenological metrics was computed by varying the threshold (normally distributed) according to the satellite product error along the modeled NDVI curve.

4.3 Results

4.3.1 The NDVI Products at Different Scales

NDVI products across all systems showed very similar patterns at the forest site (Figure 4.4). The temporal NDVI evolution describes precisely the intra-annual patterns of a (changing) canopy structure in deciduous broadleaf forests in temperate climates. Species phenology is mainly driven by temperature (and by day length to a lesser degree) (Lange et al., 2016) and can be discriminated into two phases: a phase of active/green vegetation (leafy season) from spring to fall and a phase of vegetation in-activity (dormancy) during late autumn and winter. Two distinct phenological events precede these phases: the bud burst event, tightly coupled with leaf development and maturation, as well as the onset of leaf coloring and leaf fall in autumn.

At the start of the season around April, NDVI is increasing until it reaches a maximum in June. The NDVI slightly decreases during the vegetation period and rapidly decreases in fall around mid-September due to senescence. It reaches its minimum value in December. The signal-to-noise-ratio decreases in winter due to snow cover, low irradiance and minor vegetation cover. Sentinel-2A images were manually examined in order to identify misclassified values and excluded if corresponding shortwave pyranometer measurements showed $DDSRR < 4$. Furthermore, tests demonstrated the applicability of BISE for the removal of these data points. NDVI observations falsely labeled as ‘clear’ by the processing software (and hence exhibiting lower values compared to a true clear pixel within a vegetation period) could be removed by the dynamic filtering procedure.

Green-up in 2015 and the peak of the vegetation period could be clearly identified by both MODIS NDVI and ground-based datasets. $NDVI_{multi}$, $NDVI_{hyper}$, and MODIS NDVI datasets showed the anticipated decrease during summer, where leaf color usually changes from light to dark green. Sentinel-2A data of 2015 were in line with $NDVI_{multi}$. Fall of 2015 was characterized by unstable illumination conditions (clouds, rain events, snow; see Figure 4.5a),

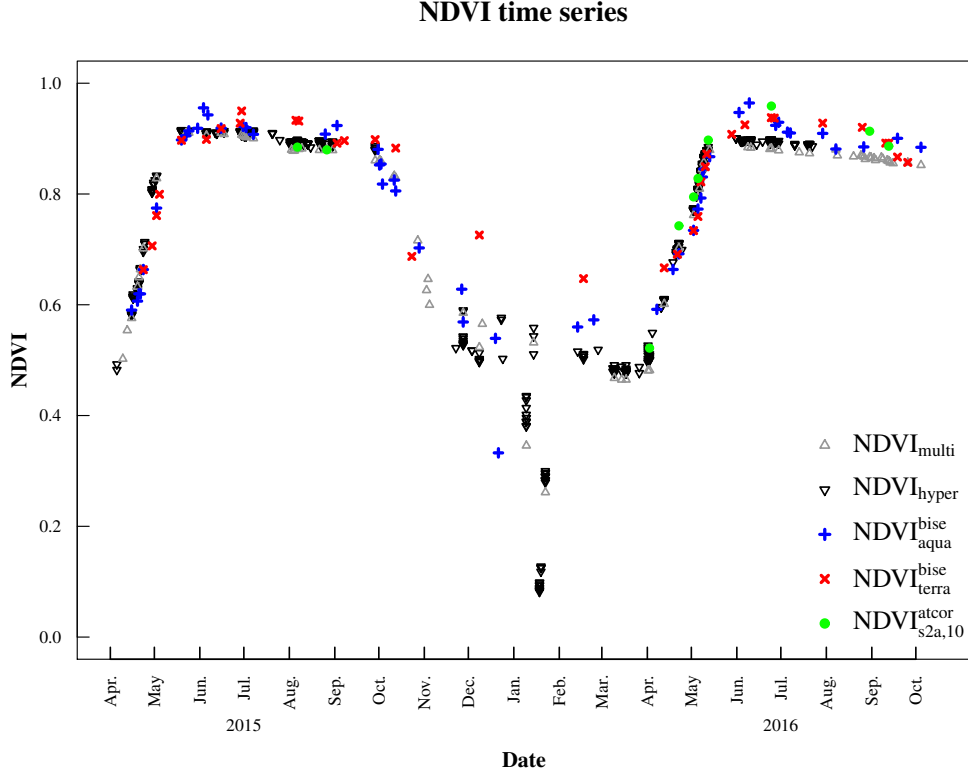


Figure 4.4: NDVI time series derived from the hyperspectral sensor system ($NDVI_{hyper}$, black triangles), multispectral sensor system ($NDVI_{multi}$, grey triangles), MODIS Aqua ($NDVI_{aqua}^{bise}$, blue plus signs) and Terra ($NDVI_{terra}^{bise}$, red crosses), as well as Sentinel-2A (10 m, Bands 4 and 8) processed with ATCOR 2/3 ($NDVI_{s2a,10}^{atcor}$, green filled circles).

which led to large gaps in satellite and hyperspectral data. In winter 2015/2016, $NDVI_{terra}^{bise}$ showed slightly higher values than $NDVI_{multi}$ and $NDVI_{hyper}$. NDVI curves exhibited a strong NDVI decrease in December 2015 and January 2016. This period was characterized by temperatures below 0 °C and precipitation events, with snow/ice cover as the main factor of decreased NDVI values (Figure 4.5b-d, confirmed by respective temperature and precipitation measurements at the tower site, which are not shown here). Data from this period were excluded from further analysis.

The initial NDVI increase in 2016 (March) was captured by all systems almost simultaneously. $NDVI_{multi}$, $NDVI_{hyper}$ and $NDVI_{s2a,10}^{atcor}$ started with lower NDVI values than $NDVI_{aqua}^{bise}$ and $NDVI_{terra}^{bise}$. NDVI evolution over the course of leaf unfolding/leaf expansion between April and May was tracked similarly by all employed sensors. Peak NDVI was reached in June, with $NDVI_{aqua}^{bise}$ showing the highest and $NDVI_{multi}$ showing the lowest values. The

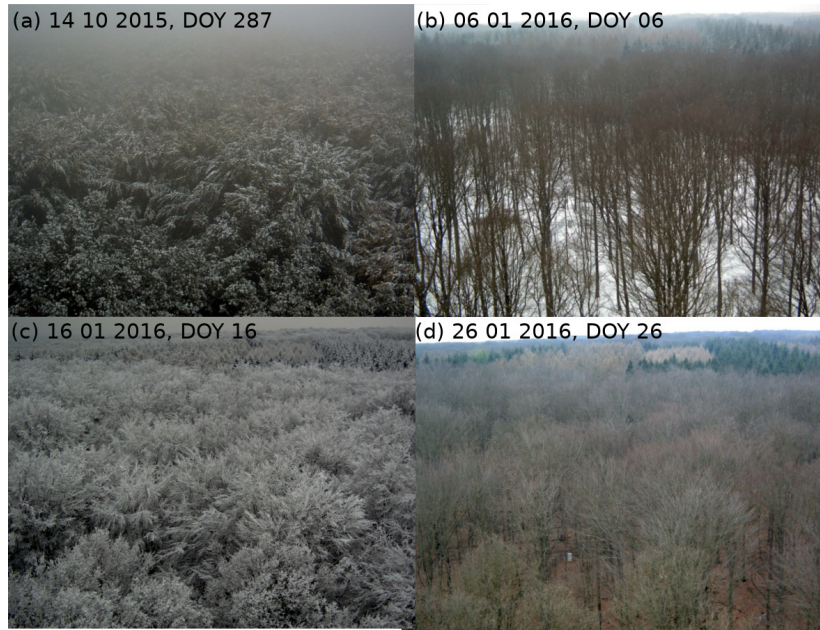


Figure 4.5: Pictures of the area around the eddy flux tower ‘Hohes Holz’. (a–d) are captured at different dates, ranging from 14 October 2015–26 January 2016. (a) shows the canopy covered by snow and ice in fall 2015. (b) shows snow cover of soil seen through a leafless tree canopy. In (c), the tree canopy is covered by ice, while (d) shows the area after a melting event (no snow or ice left).

anticipated slight decrease over the summer months was detected by all sensor systems.

Temporal mismatches between sensor systems reduced the amount of available data points for inter-sensor comparisons (Figure 4.4 versus Figures 4.6 and 4.7). NDVI products from ground spectral measurements correlated well ($r^2 = 0.998$ and $RMSE = 0.01$; Figure 4.6) and showed very similar patterns (Figure 4.4). Correlations of both ground sensors with satellite NDVI products are similar (Table 4.2) and strong ($0.72 \leq r^2 \leq 0.97$), especially with MODIS Aqua and Sentinel-2A NDVI (Figure 4.7). The magnitude of $NDVI_{multi}$ and $NDVI_{hyper}$ was comparable with satellite NDVI data points. There were some minor deviations during the vegetation period and larger deviations in fall 2015 and winter 2015/2016 (Figure 4.4). $NDVI_{multi}$ and $NDVI_{hyper}$ mainly agreed with $NDVI_{aqua}^{bise}$, showing a steady NDVI decrease towards an NDVI of around 0.56, while $NDVI_{terra}^{bise}$ showed values greater than 0.7 after November 2015. They are also apparent when looking at the correlations between ground and satellite observations (Figure 4.7).

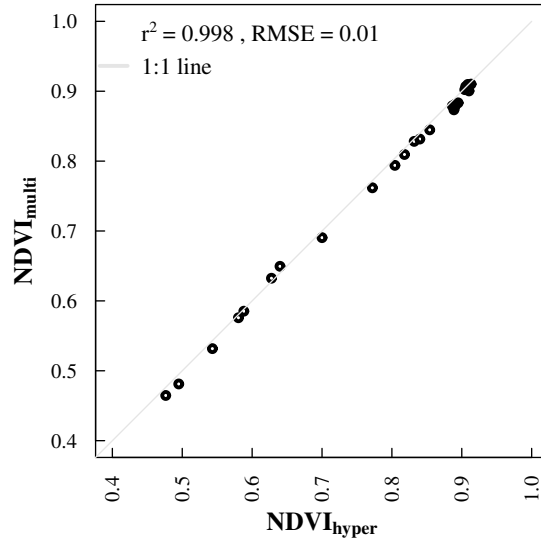


Figure 4.6: Scatter-plot of NDVI derived from the hyperspectral sensor system ($NDVI_{hyper}$) vs. NDVI from the multispectral sensor system ($NDVI_{multi}$).

	Multispectral	MODIS Aqua	MODIS Terra	Sentinel-2A
Hyperspectral	0.998	0.96	0.72	0.97
Multispectral	1	0.97	0.72	0.97
MODIS Aqua		1	0.77	-

Table 4.2: Squared Pearson correlation coefficients of NDVI product combinations from different sensors for the forest site ‘Hohes Holz’.

4.3.2 Analysis of Phenological Metrics

Phenological metrics calculated from NDVI time series differed between the sensor systems (Table 4.3), although the phenological patterns were similar. Green-up dates (GU) in 2015 varied between 25 April (DOY = 115) and 9 May (DOY = 129). GU calculated from $NDVI_{hyper}$ and $NDVI_{multi}$ were similar with standard deviations lower than one week. GU calculated from MODIS data were 8–14 days later than GU of in situ and Sentinel-2A data. Calculated green-up dates agreed with observed green-up around the end of April at ‘Hohes Holz’ (Figure 4.8a–c). Minimum NDVI in 2015 varied between 0.52 and 0.66, depending on the sensor system, and occurred between 5 April (DOY = 95) and 23 April (DOY = 113). Maximum NDVI varied between 0.92 and 0.96 and occurred between 24 May (DOY = 144)

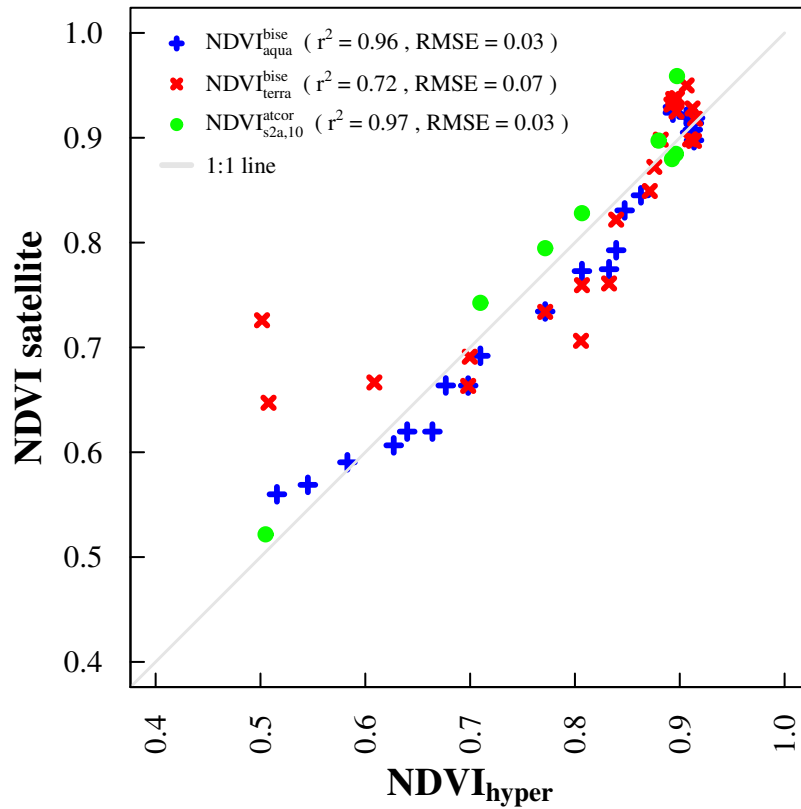


Figure 4.7: Scatter-plot of NDVI derived from the hyperspectral sensor system ($NDVI_{hyper}$) and satellite NDVI products ($NDVI_{satellite}$). Satellite NDVI products include $NDVI_{aqua}^{bise}$ (blue), $NDVI_{terra}^{bise}$ (red), as well as $NDVI_{s2a,10}^{atcor}$ (green). $NDVI_{multi}$ and $NDVI_{s2a,20}^{atcor}$ are not shown due to large similarities with $NDVI_{hyper}$ and $NDVI_{s2a,10}^{atcor}$, respectively.

and 29 June (DOY = 180). Senescence (SEN) occurred between 13 October (DOY = 286) and 24 October (DOY = 297) and showed a standard deviation of 4.3–12.0 days. SEN calculated from MODIS data occurred around 4–11 days earlier than SEN calculated from in situ and Sentinel-2A data. Calculated senescence dates mainly agreed with observed senescence around the end of October at ‘Hohes Holz’ (Figure 4.8d–f).

GU varied between 23 April (DOY = 114) and 8 May (DOY = 129) in 2016 with standard deviations lower than one week. GU calculated from $NDVI_{hyper}$ and $NDVI_{multi}$ differed by approximately one day. GU from $NDVI_{s2a,10}^{atcor}$ occurred 5–6 days later, whereas GU from MODIS data occurred 13–16 days later. Calculated green-up dates agreed with observed phenology at ‘Hohes Holz’ (Figure 4.8g–i). Minimum NDVI varied between 0.49 and 0.65

4 Monitoring and Validation of Phenology Products

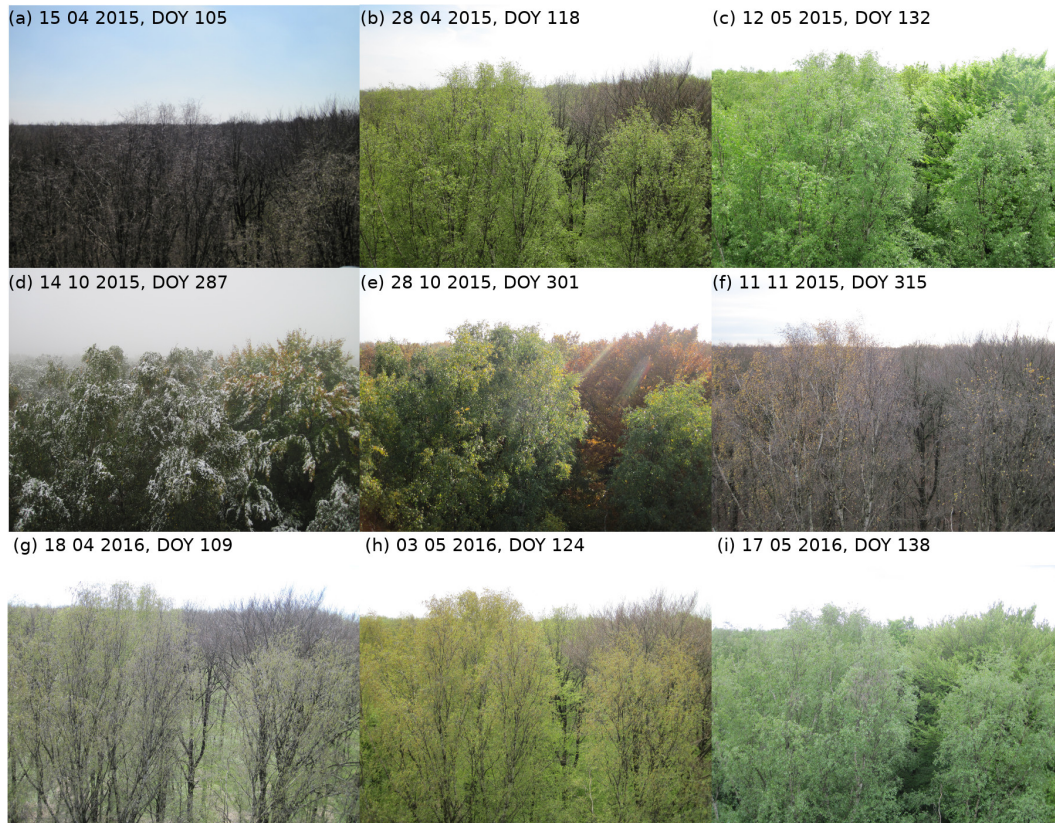


Figure 4.8: Pictures of phenological state of vegetation around the eddy flux tower ‘Hohes Holz’. (a–i) are captured at different dates, ranging from 15 April 2015–17 June 2016. (a–c) show the green-up period in 2015, (d–f) the senescence period of 2015 and (g–i) the green-up period of 2016.

and occurred between 13 February (DOY = 44) and 2 April (DOY = 93). The latter minimum NDVI was the first observation with clear conditions of Sentinel-2A in 2016. Maximum NDVI varied between 0.89 and 0.96 and occurred between 31 May (DOY = 152) and 24 June (DOY = 176). Calculated phenological metrics mainly agreed with each other and showed standard deviations consistent with the natural intra- and inter-annual phenological variability in the order of several weeks (Schaber and Badeck, 2002; Doktor, 2008).

4.4 Discussion

We validated different satellite phenology products by comparing them to multi-sensor ground measurements. The approach relates to Land Product Validation Stage 1 of CEOS

	Hyperspectral	Multispectral	MODIS Aqua	MODIS Terra	Sentinel-2A
2015/2016					
DOY Green-up	115/115	115/114	123/127	129/129	-/118
sd(Green-up)	2.1/5.2	2.4/4.3	3.8/3.2	6.4/3.3	-/5.2
DOY min(NDVI)	95/77	99/74	105/44	113/48	-/93
DOY max(NDVI)	166/152	144/160	154/161	180/176	-/176
min(NDVI)	0.49/0.49	0.50/0.46	0.54/0.56	0.66/0.65	-/0.52
max(NDVI)	0.92/0.90	0.91/0.88	0.96/0.96	0.95/0.94	-/0.96
DOY Senescence	295/-	297/-	291/-	286/-	-/-
sd(Senescence)	5.4/-	4.2/-	6.8/-	12.4/-	-/-

Table 4.3: Phenological metrics extracted from the NDVI time series with the R-package ‘phenex’. These metrics include the Day Of the Year (DOY) of green-up, senescence and respective standard deviations (sd), minimum and maximum NDVI, as well as the DOY of their occurrences. A dash (-) indicates if the metric could not be extracted from the specific time series due to unavailable data.

(Committee on Earth Observation Satellites) (Morisette et al., 2006). This required the establishment of a validation site equipped with unattended multispectral and hyperspectral sensor systems for continuous vegetation monitoring. The DFOV systems enable reflectance factor acquisition with (near-)simultaneous measurements of downwelling irradiance and upwelling radiances at high frequency. The approach chosen in this study with a single-sensor DFOV hyperspectral system reduces the costs and effort compared to a system with two spectrometers since an additional permanently-installed spectrometer is not needed (Porcar-Castell et al., 2015). Considerable instability in the scaling of the hyperspectral reflectance factor was observed when comparing time series of hyperspectral with multispectral observations. However, the scaling factor was observed to be independent of wavelength and thus does not affect NDVI, as wavelength independent scaling factors cancel out. For other applications requiring correct scaling of the hyperspectral reflectance factor, a further correction of the calibration instabilities is required. An approach to achieve this will be described in a separate publication. It has also been shown that multispectral data accuracy can potentially be improved by conducting more frequent in situ calibration/validation measurements (Jin and Eklundh, 2015).

The high correlation of NDVI time series between sensors suggests that the multispectral sensor system is sufficient for phenological pattern analysis. The additional use of a hyperspectral system provides, however, information beneficial for the validation of other remote sensing products (Porcar-Castell et al., 2015) such as chlorophyll content and leaf area in-

dex (Doktor et al., 2014), as well as the possibility to model, for example, plant productivity (Hilker et al., 2008a; Gamon, 2015). Multispectral information (of different central wavelengths and FWHM) can further be generated from hyperspectral signals so that the latter consequently enables the validation of different optical satellite missions. Upcoming hyperspectral satellite missions, e.g., EnMap (Guanter et al., 2015), and airborne hyperspectral campaigns will benefit largely from validation sites equipped with spectrometer systems.

Differences in NDVI products are either introduced from sensor specifications (e.g., band configuration), angular effects, measurement scales (and therefore, different observed vegetation patches), calibration accuracy of satellite sensors and atmospheric correction, including cloud-detection (Vermote and Kotchenova, 2008b). While sensor specifications are mainly comparable, scales and observed vegetation patches differed significantly. Due to the MODIS coarse resolution of 250 m, analysis is limited to only one pixel, whereas we could statistically analyze a cluster (within 30 m around the eddy flux tower) of Sentinel-2A pixels with 10-m resolution for this study. Hence, the observed area had a size of around 62,500 m² (MODIS) and around 3600 m² (Sentinel-2A), respectively, compared to the observed ground area of around 72 m². We observed varying performance of Sentinel-2 processors especially in the presence of small cloud patches, cloud shadows and haze. These problems may be enhanced for MODIS pixels due to their coarser resolution and subsequently more difficulties in the detection of small-scale cloud cover. The angular configuration is different for the sensor systems: MODIS Terra and Aqua overflights differ in solar zenith angle (around five to eight degrees) due to their overpass times around 10:30 a.m. and 01:30 p.m., respectively. $NDVI_{hyper}$ and $NDVI_{multi}$ exhibit different illumination angles, introduced by the selection of the time period 11:00 a.m.–01:00 p.m. and DDSRR-filtering. Sentinel-2 observed our region of interest around 10:30 a.m. with a solar zenith angle similar to MODIS Terra. Hence, angular differences can be introduced by both temporal mismatches and angular configuration, which are difficult to differentiate. We mainly observed small NDVI differences within one day between sensor systems, although MODIS Terra showed larger differences to the other sensors during winter 2015/2016. We found considerably higher NDVI values in diurnal multispectral NDVI profiles before 10:30 a.m. only during winter time, assuming potential effects of illumination geometries on the signal. Consequently, MODIS Terra and potentially Sentinel-2 are more sensitive to this effect than MODIS Aqua due to their overpass time before noon. We could not substantiate this effect for Sentinel-2 data, since no observations with clear sky conditions were available for this period. Temporal mismatches between different systems increase using filter algorithms due to the selection of the ‘best’ or ‘true’

NDVI values. Remaining NDVI values after dynamic filtering are used as input for models restoring the phenological profile. Hence, the detection of phenological local extrema, e.g., date and magnitude of the peak of the growing season, is critical to robustly depict phenological metrics. Despite the above described effects, we could not detect a considerable influence of temporal mismatches within one day on phenological metrics extraction. This might be different for other vegetation types or climates with modified phenological timing. Temporal mismatches of several days can in extreme cases cause a derivation of potentially wrong phenological phases, e.g., Sentinel-2A observing clouds during green-up, while MODIS with higher repetition rate observes clear sky conditions in between.

In situ and satellite-based NDVI time series captured similar phenological patterns despite the numerous influencing factors. Correlations between the sensor systems were strong, although some deviations occurred during fall and winter between NDVI of ground sensor systems and MODIS Aqua and Terra. RMSEs between satellite and in situ data exhibited the same magnitude as found in another recent study (Geng et al., 2016) and as the MODIS NDVI uncertainty ($0.02 + 0.02 \times NDVI$, Vermote and Kotchenova (2008a)).

Depicted phenological metrics from both in situ datasets were consistent with each other and agreed with Sentinel-2A-derived metrics in 2016. Standard deviations and even the differences of green-up and senescence between the sensor systems were consistent with the natural variance of phenology. MODIS NDVI values during spring and fall were lower compared to in situ data, leading to a deviation in depicted phenological phases: green-up occurred later and senescence earlier. The later GU is contrary to the results found in Balzarolo et al. (2016). Tests with NDVI from hyperspectral data resampled to MODIS bands demonstrated a small shift towards later GU. A larger shift was introduced when averaging NDVI of all Sentinel-2 pixels within the MODIS pixel, although the area is characterized by a homogeneous deciduous broadleaf forest with small clearings. Since satellite observations integrate phenological signals from different species within the pixel, the observed effect of spatial scale mainly stems from different species composition. The species' signals might also be differently weighted according to the sensors' point spread function. This hypothesis is supported by high satellite NDVI values in summer, later green-up and day of maximum NDVI, as well as earlier senescence with increasing scale.

Sentinel-2A images are currently available as Level 1C products. Atmospheric, cirrus and terrain corrections are necessary for further analysis. The processing tool ATCOR 2/3 produced reasonable Level 2A products. Although ATCOR 2/3 provided scene classification

information supporting the elimination of outliers, misclassifications compromised the resulting NDVI time series. Three approaches can be distinguished eliminating faulty records in time series: images can be manually examined, but this is not feasible for long time series or large datasets. The usage of diffuse shortwave pyranometer measurements as selection criteria is also limited to observation sites with respective infrastructure, but these measurements are getting more and more common. Dynamic filter algorithms, such as BISE, reduce the risk of including false observations while being applicable to large datasets without the need for auxiliary site-specific measurements. Here, BISE was able to consistently remove faulty NDVI values. Daily MODIS NDVI values of 250 m lack scene classification information and respective analysis consequently requires the use of filter algorithms. BISE was able to restore the NDVI profile, although we detected mismatches with ground data during fall 2015 and winter 2015/2016. During December 2015 and January 2016, low temperatures, precipitation events and subsequently snow cover and frozen surfaces led to a strong decrease in multispectral and hyperspectral, as well as MODIS Aqua NDVI not related to vegetation. A strong increase in NDVI during the snow-melt period was detected, which is consistent with other studies observing NDVI of snow-covered vegetation (Jönsson et al., 2010; Eklundh et al., 2011). The length of this situation exceeded the sliding period of BISE parametrization. Consequently, decreased MODIS Aqua NDVI values were not eliminated. In contrast, MODIS Terra NDVI increased during winter 2015/2016, relating to the previously-mentioned angular effects.

Apart from assessing the accuracy and precision of satellite products, a validation site offers the opportunity to examine data gaps or periods with low signal-to-noise-ratio introduced by bad illumination conditions and cloud or snow cover. Hence, complete ground NDVI time series are beneficial. Future studies may use a combination of DDSRR-filtering and dynamic filtering, subsequently lowering the DDSRR-threshold to obtain more data points under overcast or hazy conditions, increasing the number of usable ground measurements. The comparison of filtered satellite-based time series with in situ data provides the possibility to modify BISE parameters for satellite phenology depiction, if necessary (Xu et al., 2017).

4.5 Conclusions

Here, we validated satellite phenology products for a deciduous broadleaf forest. Phenological metrics were consistent between sensor systems and agreed with the natural phenological

variance also with sparse data coverage during autumn and winter time. Methods taking the effects of strongly diffuse illumination into account were already addressed in other studies (Hilker et al., 2008b; Pacheco-Labrador et al., 2016). In our measurements, however, we mostly excluded periods of time with highly diffuse illumination because of unstable measurements. A larger number of measurements under highly diffuse and low illumination conditions could be retained by continuous integration time optimization (low illumination) (Sakowska et al., 2015; Pacheco-Labrador et al., 2016) or averaging over a larger number of consequent measurements (diffuse conditions). Further, spectral sensors with technical specifications such as research-grade instruments of the WMO (World Meteorological Organization) class quality are beneficial for dealing with these suboptimal conditions. Validation sites should utilize either multispectral sensors with bands chosen according to satellite sensors' channels or spectrometers simulating respective bands, since the band configuration has an effect on extracted phenological metrics. Improvements in observing the 'true' phenological signal in phenological time series were detected by the use of Sentinel-2A data compared to MODIS. Nevertheless, despite the long service and comparably coarse spatial resolution of MODIS Aqua and Terra, MODIS-based phenological metrics agreed fairly well with ground based products. The high spatial resolution of Sentinel-2 reduces mixed-pixel effects ubiquitous in phenology research and offers new opportunities to understand connections between site measurements and large-scale processes, e.g., in the analysis of eddy covariance footprints or vegetation-dependent processes in heterogeneous landscapes. Temporal mismatches between phenological phases and timing of Sentinel-2 observations, under certain conditions critical for the extraction of phenological metrics, will decrease with the upcoming availability of Sentinel-2B data and hence increased temporal resolution of Sentinel-2. The Sentinel-2 processing tool ATCOR 2/3 allows for atmospheric, terrain and cirrus correction of Sentinel-2 images and provides scene classification information on cloud cover, illumination condition and saturation. Despite these comprehensive ancillary data, dynamic filtering algorithms like BISE are still beneficial to generate reliable phenological time series.

5

Continuous calibration monitoring of automated hyperspectral measurements of vegetation canopy optical properties. I. Radiometric and reflectance calibration

Authors Benjamin Dechant^{1,2,3,*}, Maximilian Lange¹, Matthias Cuntz^{2,4} and Daniel Doktor¹

Affiliations ¹ Department Computational Landscape Ecology, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany
² Department Computational Hydrosystems, Helmholtz-Centre for Environmental Research-UFZ, Permoserstr. 15, 04318 Leipzig, Germany
³ Now at Research Institute of Agriculture and Life Science, Seoul National University, Republic of Korea
⁴ INRA, Université de Lorraine, AgroParisTech, UMR Silva, 54000 Nancy, France
* Correspondence: benjamin.dechant@ufz.de; Tel.: +49-(0)341-235-1010

In prep. for: Remote Sensing

State: to be submitted

Acknowledgements We thank L. Dienstbach, G. Schuhmann, H. Zöphel and I. Quirós for supporting us with the field activities, H. Hertel for the Labview programming of the spectrometer and H. Zöphel for the logger programming and support in the field. We are particularly grateful for the assistance of S. Gimper who helped with the installation and maintenance of the sensors, various technical issues and assisted in field measurements. Without the support of C. Rebmann, who is PI for the eddy covariance sites, our work would not have been possible. We thank Niklas Hase for providing critical feedback on remaining calibration issues in earlier stages of the analyses. We also acknowledge helpful technical advice from Alasdair MacArthur and other researchers in the framework of the COST actions EUROSPEC and OPTIMISE.

Author Contributions B.D., M.C. and D.D. conceived and designed the experiments; B.D. performed the experiments and analyzed the data; M.L. assisted with installation and maintenance of the sensors as well as the experiments and contributed to data processing; B.D. wrote the paper, all co-authors edited the manuscript.

Abstract Automated hyperspectral measurements in the optical and near-infrared range are increasingly used to complement eddy covariance measurements and provide data for ground validation of air- and spaceborne observations. The calibration of such hyperspectral measurements is crucial but typically only assessed on a campaign basis rather than continuously. We therefore examined the long-term stability of the absolute radiometric and reflectance calibrations of automated measurement systems for hyperspectral vegetation reflectance on a continuous basis using more stable multispectral observations as reference. The study was conducted simultaneously at two sites with almost identical automated hyper- and multispectral systems but differing in type of vegetation: grassland vs. temperate broadleaf forest. The commonly used campaign-based method using white reference observations was also applied for comparison at the grassland site. We observed considerable instability in the absolute reflectance calibration of both hyperspectral systems. The observations of both up- and downwelling radiation fluxes contributed to the instabilities. The latter were found to consist of incorrect time-dependent but wavelength-independent scaling of the reflectance spectra that included both a large discontinuity at the grassland site as well as more gradual drifts at both sites. Especially the latter could only be corrected reliably with the help of the continuous multispectral observations. This led to a considerable decrease in errors by factors varying between two and seven such that an agreement between daily aver-

age multi- and hyperspectral reflectance observations at a level of less than 10% at the forest and less than 15% at the grassland site could be achieved. The agreement after correction in terms of squared Pearson correlation was larger or equal to 0.9 in all cases. Our results show the need for continuous calibration monitoring and correction of automated hyperspectral measurements and suggest a method to achieve this at relatively low cost.

5.1 Introduction

Automated optical measurement systems are increasingly used at eddy flux tower sites (Balzarolo et al., 2011; Eklundh et al., 2011; Porcar-Castell et al., 2015). The motivation is typically to study the feasibility of using optical methods to estimate carbon fluxes, mostly Gross Primary Productivity (GPP) (Balzarolo et al., 2015; Hilker et al., 2010; Nakaji et al., 2007; Rossini et al., 2010, 2012; Sakowska et al., 2014; Yang et al., 2015, 2017), as well as to continuously monitor vegetation phenology and stress by estimating the relevant leaf and canopy traits (Balzarolo et al., 2016; Sakowska et al., 2016; Soudani et al., 2012; Wingate et al., 2015). The estimation of transpiration and heat fluxes from optical measurements have also been investigated (Hilker et al., 2013) and optical measurements were combined with eddy covariance observations to investigate daytime ecosystem respiration (Hilker et al., 2014). Furthermore, the optical observations can be used for continuous ground validation of airborne and satellite measurements, which often motivates the choice of measurement instrument and its configuration (Eklundh et al., 2011; Lange et al., 2017).

Broadband observations of long and short wave incoming and reflected radiation are routinely measured at eddy covariance sites but more detailed, spectrally resolved measurements of incoming and reflected light are only measured at some of the sites (Balzarolo et al., 2011). Multispectral sensors that can be used to calculate vegetation indices involving only a few spectral bands such as the Normalised Difference Vegetation Index (NDVI) (Tucker, 1979) are now widely installed at eddy covariance sites (Balzarolo et al., 2011; Eklundh et al., 2011). Such multispectral sensors based on bandpass filters are typically used to measure both incoming and reflected sunlight simultaneously by having an upward and a downward looking sensor, which results in calibrated values of the reflectance factor (Jin and Eklundh, 2015). These instruments tend to require relatively little effort to install and maintain (Porcar-Castell et al., 2015) as the observations can be recorded with commonly used data loggers and typically no fiber optic cables are necessary as the detectors are located in the same hous-

ing as the multispectral filters. The radiometric calibration of these instruments is normally conducted by the manufacturers but validation of the absolute reflectance calibration with a white reference (WR) panel is recommended (Jin and Eklundh, 2015).

Hyperspectral observations consisting of a large number of narrow spectral bands can be used both for more flexible choice of wavelengths for vegetation indices as well as applications such as radiative transfer model inversion (Jacquemoud, 1993) and the study of sun-induced chlorophyll fluorescence (Meroni et al., 2009; Porcar-Castell et al., 2014). The spectrometer-based systems required to conduct such hyperspectral observations, however, require more careful calibration and calibration monitoring (Balzarolo et al., 2011; Pacheco-Labrador and Martín, 2015; Porcar-Castell et al., 2015). This and the significantly higher costs are the reason why only a relatively small number of such systems have been used and reported in the literature (Cogliati et al., 2015; Daumard et al., 2010; Drolet et al., 2014; Gamon et al., 2006a; Hilker et al., 2007; Huber et al., 2014; Leuning et al., 2006; Meroni et al., 2011; Nakaji et al., 2007; Sakowska et al., 2015; Yang et al., 2015). Typically, the spectrometers are connected to the point of measurement with fiber optic cables that are prone to signal losses and are one factor why unattended hyperspectral observations tend to be less stable than filter-based multispectral measurements. Similar to the set-up for multispectral measurements, two spectrometers with upward and downward looking fields-of-view could be used. This, however, would imply buying two rather expensive instruments as well as conducting a cross-calibration of the spectrometers. Alternatively, a single spectrometer can also be used in dual-field-of-view (DFOV) mode (Porcar-Castell et al., 2015) with a mechanism that enables to measure both incoming and reflected sunlight successively with a relatively short time delay. Commercial portable field spectrometers with rather stable operation have been used at some field sites (Huber et al., 2014; Sakowska et al., 2015) but most research groups have developed their own automated spectrometer systems on the basis of commercial components (Hilker et al., 2007; Leuning et al., 2006; Meroni et al., 2011; Tortini et al., 2015). Particularly for the latter type of system, a thorough assessment of calibration quality over time is required.

Spectrometers typically require two types of calibration: wavelength calibration ensuring that the position of the measurement channels in the electromagnetic spectrum is well characterized and the radiometric calibration ensuring that the detector counts are converted correctly into physical units. For hyperspectral measurements of vegetation reflectance, wavelength calibration is crucial, while radiometric calibration is not required in the absolute sense for the upward and downward looking sensors separately. This, however, is not true for certain

applications such as measurements of sun-induced chlorophyll fluorescence, which require observations of both up- and downwelling radiation in absolute physical units. Model related applications such as radiative transfer model inversion for biophysical parameter estimation (Jacquemoud, 1993) typically require absolute reflectance calibration, which is also necessary for the purpose of satellite reflectance product validation. Reflectance ratio or normalized difference indices, in contrast, are not affected by a simple scaling of the whole reflectance spectrum as the scaling factor cancels out.

Absolute reflectance calibration is typically performed by making a reference measurement with a white reference (WR) panel on a campaign basis (Balzarolo et al., 2011). Only few automated systems conduct reference measurements simultaneous to or before each reflectance measurement (Huber et al., 2014; Sakowska et al., 2015). Instead of WR panels, reference light sources are also used for campaign-based absolute reflectance and radiometric calibrations and cross-calibration to a reference spectrometer was also reported in the literature (Balzarolo et al., 2011; Meroni et al., 2011). The problem for both radiometric and wavelength calibration on a campaign basis is that even if instabilities in the calibrations are identified, they cannot be corrected continuously. The latter may, however, be necessary to obtain reliable data in the case of potentially nonlinear drift. Moreover, some calibration instabilities occurring between calibration campaigns may not be detected at all and therefore also cannot be corrected. The need for further investigation of the temporal (and thermal stability) of spectrometer calibration for long term in-situ measurements was also identified in an article describing the activities and findings of the EUROSPEC Cost Action by Porcar-Castell et al. (2015).

Correct radiometric calibration of both up- and downwelling observations implies accurate reflectance calibration. However, reflectance is much more sensitive to calibration errors as it does not contain the main driving signal of up- and downwelling radiation fluxes, namely the incoming solar irradiance. This together with the importance of reflectance observations for vegetation studies motivates the emphasis of reflectance in this study but the separate radiometric calibrations are taken into account for a better understanding of the underlying causes of reflectance calibration instabilities.

The objective of this study was 1) the characterization of the long-term temporal stability of the reflectance and radiometric calibrations of two automated hyperspectral systems using a continuous monitoring approach and 2) the demonstration of an approach for continuous correction of calibration instabilities and drift. Our main hypotheses were 1) that the hy-

perspectral reflectance and radiometric calibrations can show considerable instabilities on seasonal time scales, 2) that the main effect is a wavelength independent scaling factor, and 3) that multispectral observations are more stable and can therefore be used to continuously correct instabilities in the hyperspectral observations.

5.2 Materials and Methods

5.2.1 Study sites

The measurements were conducted at two sites in central Germany, both equipped with eddy covariance systems to monitor carbon and water fluxes as well as a number of additional meteorological instruments.

Grassland The first site is a wet grassland site called ‘Grosser Bruch’ (52° 1’47.31"N, 11° 6’17.45"E), which is characterized by a high water table and intermittent grazing by cows as well as occasional harvesting of grass. All measurements are conducted in an enclosure, which is managed in terms of cutting the grass in order to better represent the conditions outside of the enclosure in the eddy flux tower footprint. The height of the eddy tower where sensors are installed is about 2 m. This site was the focus for campaign-based calibration measurements due to the easier accessibility of the sensors.

Hohes Holz The second site, ‘Hohes Holz’, is located in a temperate broadleaf forest at a distance of about 10 km from the grassland site. The eddy flux tower is 50 m high and about 20 m above the tree canopy. The site was already more comprehensively described in Lange et al. (2017).

5.2.2 Measurement set-up

We installed automated hyper- and multispectral measurement systems with very similar instruments and configurations at the two sites. At each site, the hyper- and multispectral systems have quasi-identical field-of-views (FOV) and the spectral ranges of the hyperspectral instruments fully cover the multispectral channels. The hyperspectral and multispectral

systems were installed in summer 2014 at the grassland and in winter 2014 at the forest site and available data until fall 2016 was analyzed.

Hyperspectral systems Both hyperspectral systems are dual-field-of-view (DFOV) type systems with a single spectrometer, a bifurcated optical fiber and shutters to measure hemispherical-conical reflectance factors. The details of the optical fibers and shutters, the software as well as a comprehensive description of the system at the forest site is given in Lange et al. (2017). The systems at the two sites differ almost exclusively in the sensor types but not the set-up and measuring protocols. An exception to this is the viewing angle of downward looking sensors, which is off-nadir at the forest site but at nadir at the grassland site. For the hyperspectral system, a QE65PRO spectrometer (Ocean Optics, Dunedin, FL, USA) was used at the grassland site instead of the older QE65000 at the forest site. The spectrometer at the grassland site measured in the range of 400-1180 nm but the range was restricted to 440-1100 nm to avoid excessive noise at the edges of the range. Analogously, the range at the forest site was restricted from 400-1170 nm to 440-1010 nm. The nominal spectral resolution was 1.6 nm at the grassland site and a similar value was observed at the forest site. All observations were corrected for the measured dark current. The hyperspectral measurements were collected every 10 minutes with three repeated cycles of dark current measurement, upward and downward looking measurements, and dark current measurement. These measurements had typical integration times between two and four seconds.

Multispectral systems For the multispectral system at the grassland site, SKYE SKR 1860 multispectral sensors (Skye Instruments Ltd., Llandrindod Wells, Powys, U.K.) equipped with four spectral channels and weather-proof housing were installed instead of the older SKR 1850 model at the forest site. The exact optical specifications for both systems are given in Table 5.1. The multispectral sensors were radiometrically calibrated by the manufacturer prior to installation. The FOV configuration of the multispectral sensors was very similar to the one for the hyperspectral system (Lange et al., 2017). The multispectral data were collected as 10 second averages for the whole period.

Channel	1	2	3	4
CW (nm)	658 (660)	708 (709)	739 (740)	858 (860)
FWHM (nm)	45 (38)	10 (10)	10 (10)	11 (11)

Table 5.1: Overview of the specifications of the multispectral channels: central wavelength (CW) and full width at half maximum (FWHM) in nanometers. The values are given for the grassland site and the forest site in parentheses. The mean values of upward and downward looking sensors are reported for each site, as their respective specifications were almost identical.

5.2.3 Calibration, monitoring and correction of hyperspectral systems

Overview of calibration procedures Several steps were necessary for the calibration of the hyperspectral systems and a flowchart is presented in Fig. 5.1. At the grassland site, the system spectral response for the upward looking fiber was cross-calibrated with a field spectrometer on one day. On the same day, a white reference (WR) panel measurement was conducted to measure the ratio of the downward to upward looking fibre. The latter measurement was repeated several times in the years 2015 and 2016 at the grassland site to monitor stability. The one-time absolute reflectance as well as radiometric calibrations of up- and downwelling observations were then achieved by using the multispectral observations to scale the spectra on one day on the basis of daily average values. This step was necessary as the WR measurements appeared not to be scaled correctly potentially due to partial obstruction of the sky above the WR panel by the mast arm structures and sensors installed on it or a small contribution of vegetation upwelling radiance reflected off the mast arm structures (see section 5.2.3). The relative spectral calibration, i.e. the spectral shape, is assumed to be mostly unaffected by this. Finally, a continuous calibration correction was conducted using the daily average ratios of multi- and resampled hyperspectral reflectance and compared to the one-time calibration to assess temporal stability. The resampling was done by using the hyperspectral observations calibrated in terms of spectral shape with the first two steps described above and then convolving with the spectral response curves of the multispectral sensors provided by the manufacturer. For convenience, the hyperspectral observations resampled to the multispectral channels are labelled ‘resampled hyperspectral channels’ in the following. For all analyses involving a comparison of multi- and hyperspectral observations, the temporally closest multispectral observations for each hyperspectral measurement were

selected. More details on individual calibration steps are given in the sections below.

The calibration procedure for the forest site followed the same logic but minor differences are described in more detail in the sections below.

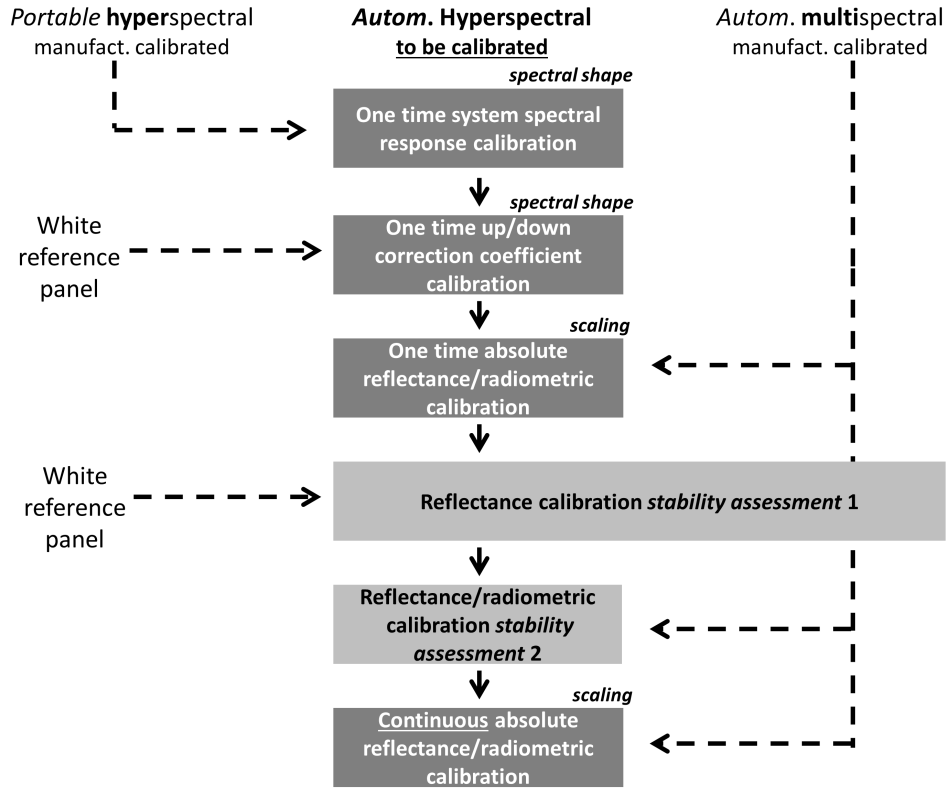


Figure 5.1: Overview of the different sensors, calibration and assessment steps. Automated (‘autom.’) systems are the continuously measuring systems, while ‘portable’ refers to a field spectrometer that was only used in measurement campaigns. ‘Up/down correction coefficient’ refers to the ratio of the upwelling and downwelling radiation fluxes when measuring the same signal with the help of the white reference panel. ‘Spectral shape’ indicates a wavelength dependent calibration factor, while ‘scaling’ indicates a scaling factor independent of wavelength. ‘reflectance/radiometric’ indicates that the calibration could be applied to either of the two as in the case of reflectance a full radiometric calibration is not necessary. Calibration steps are shown as dark grey boxes with white letters, while assessment steps are shown as light grey boxes with black letters.

Calibration of system spectral response and up/down correction factor

The calibration of the spectral response of the hyperspectral system at the grassland site was conducted in July 2016 by using a calibrated FS4 field spectrometer (ASD, Boulder, CO,

USA). This was done by interpolating the FS4 radiance spectra of the WR panel to the measurement channels of the QE65PRO spectrometer, calculating the ratio of the interpolated FS4 radiance and the QE65PRO spectra of the upward looking fiber and smoothing the ratio using a moving average. The latter was necessary as the different spectral resolutions and a small spectral shift of the two spectrometers led to artificial features in the ratio that were not due to the system or spectrometer spectral response. This smoothing procedure would be also necessary when resampling the higher resolution spectra to the lower resolution as even very small spectral shifts that cannot be corrected anymore in practice lead to considerable peaks in the ratio. The corresponding spectral response for the downward looking fiber measurements was obtained by multiplying the upward calibration factors with the ratio of down- to upward looking observations obtained with a calibrated WR panel (SphereOptics, Uhldingen, Germany). The resulting correction spectra for the up- and downward looking fibers were then applied to all measurements of the hyperspectral systems conducted before and after the date of calibration.

The procedure at the forest site followed the same steps except that the field spectrometer observation was replaced with a calibrated grassland observation on a clear-sky day. This approach is based on the assumption that the atmospheric conditions at these two sites were comparable on that day due to the small spatial distance between them. A direct comparison of hyperspectral solar irradiance after calibration with the corresponding reference measurements is shown in Fig. 5.6 of the supplementary material.

In principle, the spectral response characterization would not be required for reflectance observations as the up/down correction coefficient would be sufficient. However, the resampling step for the comparison with multispectral observations requires at least a relative spectral response calibration and therefore this calibration step is necessary in practice for our method even if absolute radiometric calibration of up- and downwelling observations is not of interest.

Campaign-based assessment of reflectance calibration stability using a WR panel Measurements with the WR panel (SphereOptics, Uhldingen, Germany) were conducted on a campaign basis to study the temporal stability of the hyperspectral reflectance calibration with a commonly used approach. A total of five WR measurements were conducted for the hyperspectral system at the grassland site and were corrected for the reflectance of the WR panel. A small contribution of the vegetation signal was detected in some of the measurements using the hyperspectral system which were probably caused by

reflections off the eddy tower mast arm and surfaces of measurement instruments installed on the mast arm. These vegetation signals did not always have the same contribution to the signal, but the presence of the signal could easily be detected due to the clear red-edge vegetation signature (Fig. 5.7). We corrected the WR measurements by subtracting the scaled pure vegetation signal measured on the respective day from the WR radiance in such a way as to minimize the red-edge signature. This approach led to considerable improvements but could not always entirely remove the effect of the red-edge vegetation signal on the spectral shape.

The WR measurements were not repeated at the forest site due to the more difficult access to the sensors as well as the challenges to obtain high quality observations due to shading/obstruction by tower and mast arm structures.

Continuous monitoring of temporal stability of reflectance and radiometric calibrations and multispectral scaling corrections We verified that the multispectral observations were more stable due to their considerably simpler measurement set-up without optical fibers and could therefore be used to monitor both temporal and spectral stability of the hyperspectral system and apply corrections if necessary. Channel 4 (858 nm) of the multispectral upward looking sensor at the grassland site was found to yield incorrect measurements, however. The scaled observations of the neighboring channel 3 were therefore used as replacement. This approach is based on the very high correlation of solar irradiance measured in spectrally close spectral channels presented in more detail elsewhere (Dechant et al., 2018), with scaling coefficients obtained from WR measurements. A similar procedure was applied at the forest site for the faulty detector of the upward looking sensor channel 1.

We conducted the analysis both for up- and downwelling radiation fluxes separately. We investigated the data at seasonal scale as well as calculating Pearson correlation and conducting linear regressions on a daily basis to detect seasonally changing relationships. The linear regressions were subsequently reduced to simple daily average ratios to obtain the slope values as there were no indications for significant offsets. The reflectance measurements of the multi- and hyperspectral systems were compared by calculating daily averages between 10:00 and 14:00 CET. Data were filtered by multiple thresholds to exclude unstable, low quality observations due to various technical issues such as failure of the shutters switching between up and downward looking observations at the grassland site as well as snow. This unfortunately led to a considerable fraction of the available data to be excluded.

First, the hyperspectral reflectance observations were only scaled with a temporally constant correction factor on one day to correct for effects mentioned in section 5.2.3. The choice of the day for one-time calibration is necessarily somewhat arbitrary, but the evaluation using correlation measures is not affected by this. In order to also minimize the effects on error and bias measures, days within a rather stable period in spring 2015 were selected. Second, a continuous correction was conducted by using the daily average ratio of the multispectral and the corresponding resampled hyperspectral observations as spectrally constant scaling factors. The daily average ratios for each channel were averaged over the four channels to obtain a single correction factor for the whole spectrum. This method was applied to the up- and downwelling fluxes separately, although it could be applied directly to reflectances if absolute radiometric calibration of the former is not required.

5.3 Results

5.3.1 Campaign-based white reference measurements for calibration stability monitoring at the grassland site

Overall, the multispectral system at the grassland site showed much better temporal stability of the reflectance calibration than the hyperspectral system. We found that the ratio of up- over downwelling radiation fluxes for the calibrated multispectral sensor system as observed with white reference (WR) panels stayed close to the ideal value of 1 over a period of about 18 months (overall mean value 1.01, coefficient of variation $CV = 0.04$, Fig. 5.2a). The observed temporal pattern was consistent for all multispectral channels (coefficient of variation averaged over all dates, $CV = 0.02$) and a somewhat larger deviation between them only occurred on the first measurement date ($CV = 0.05$).

For the hyperspectral system, in contrast, we found large temporal variation of the up- over downwelling ratio (mean value 0.85, $CV = 0.38$), which had a minimum value of 0.55 and a maximum value of 1.38. However, the temporal pattern was very consistent for all resampled hyperspectral channels with an average CV of 0.006. We also did not observe a clear change in the spectral shape of the correction factor between upward and downward looking fibers (Fig. 5.2b). The 2016-07-26 measurement had a lower slope in the visible range than the other observations, however, and the 2015-07-22 measurement showed larger differences

in the near-infrared range as well as higher noise. Nevertheless, the deviations from the reference observation on 2016-07-26 were mostly smaller than 5% after normalizing the spectra at a wavelength of 680 nm (Fig. 5.2b).

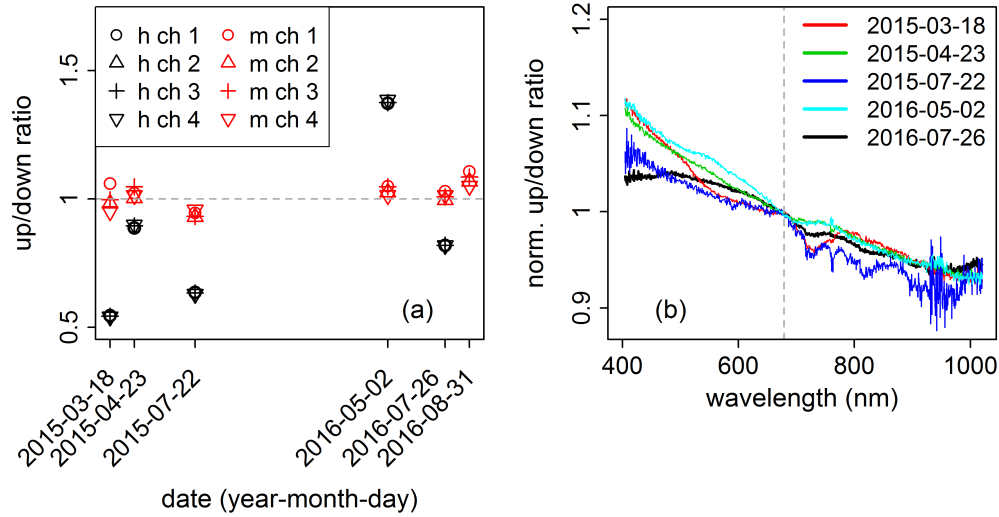


Figure 5.2: Overview of ratio of upwelling over downwelling observations for the same light conditions for the multi- and hyperspectral systems at the grassland site. A white reference (WR) panel was used for the upwelling observations. Panel (a) shows the time series of the WR measurements for the resampled hyperspectral channels 1-4 (e.g. ‘h ch 1’) after one-time calibration as black symbols and the multispectral channels 1-4 (e.g. ‘m ch 1’) as red symbols. The dashed grey horizontal line indicates the ideal value of 1. There were no hyperspectral observations for the last date. Panel (b) shows the corresponding spectral shapes of the WR observations. The measurements were normalized at 680 nm (dashed grey reference line) for easier visual comparison of the spectral shape. In order to show the full characteristics of the spectral shape of all observations in (b), the one-time calibration was not applied here.

5.3.2 Continuous stability monitoring of hyperspectral observations

Seasonal vegetation dynamics and overview of stability The grassland site showed repeating patterns of fast increase in the near-infrared reflectance (multispectral channel 4 at 858 nm) up to values as high as 0.6 and sudden decreases due to cutting of

the grass or senescence (Fig. 5.3d). The lowest values of about 0.1 were reached during the winter in early 2015. The visible reflectance (multispectral channel 1 at 658 nm) mostly showed very low values around 0.025 but increased to about 0.15 both in spring 2015 and 2016 when near-infrared reflectance sharply dropped.

The forest site was characterized by the phenology of the broadleaf trees with near-infrared reflectance varying between about 0.35 in summer and 0.15 in winter. The maximum values reached in the 2016 vegetation period were somewhat lower than in 2015. The visible reflectance varied between about 0.015 during the vegetation period and 0.07 in winter.

Only days with Pearson correlation larger than 0.8 between resampled hyper- and multispectral radiation fluxes as well as stable average daily ratios were selected for further analysis. However, the latter showed considerable temporal variation at both sites, both for down- and upwelling fluxes (Fig. 5.3).

Grassland site We found considerable temporal variation in the daily average ratios of hyper- over multispectral observations for the up- and downwelling components as well as the reflectance (Fig. 5.3a-c). The largest change was a sudden drop in the ratio of the downwelling component from a value of about 1 to 0.5, which occurred in fall 2015 (Fig. 5.3a). Both before and after this event, however, a more gradual upward drift could be observed. The upwelling ratio was overall more stable but showed also drifts of varying sign as well as a drop in the signal at about the same time as for the downwelling ratio that subsequently recovered (Fig. 5.3b). The reflectance ratios (Fig. 5.3c) showed a combination of features from the temporal patterns of both the up- and downwelling ratios. This resulted in a more stable behavior before the discontinuity in the downwelling ratio and a strong downward drift after this event. The corresponding values obtained from the WR observations (section 5.3.1) showed reasonable agreement with the overall temporal behavior as well as the actual values of the hyper- over multispectral ratios on four of the five measurement days (Fig. 5.3c). There was a large difference on the first measurement day, however, where the WR observation was considerably lower than the daily average ratio.

On closer inspection, it appears that decreases in the ratio of the upwelling component were partly associated with increases in the near-infrared reflectance (Fig. 5.3b).

Forest site Considerable temporal variation in the daily ratios of hyper- over multispectral observations were also observed at the forest site (Fig. 5.3e-g). This gradual variation

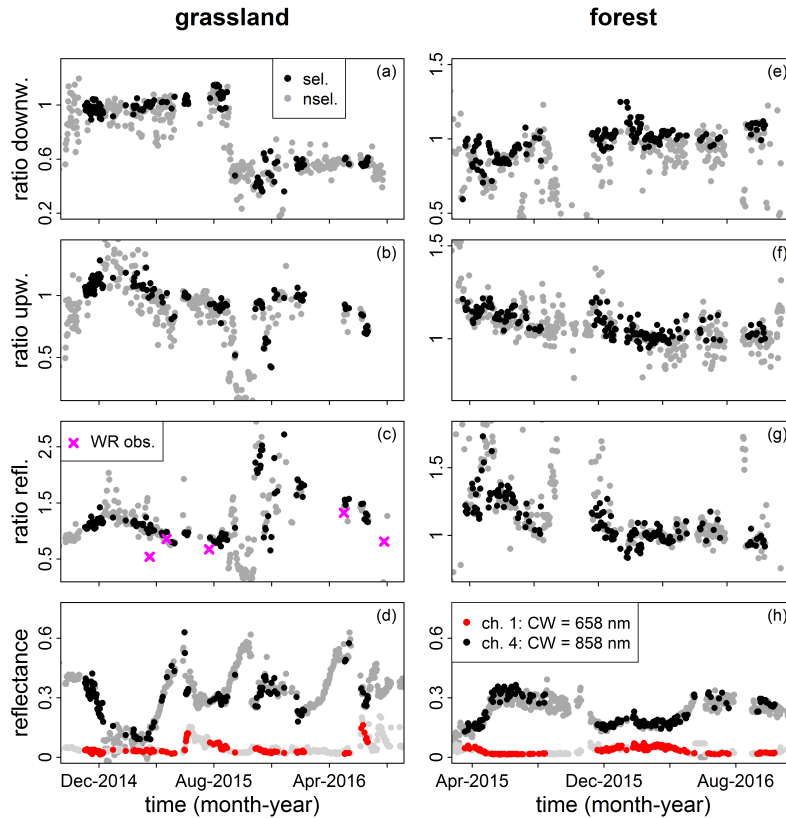


Figure 5.3: Overview of daily ratios between multi- and corresponding resampled hyperspectral observations, averaged over all four multispectral channels of (a, e) downwelling, (b, f) upwelling and (c, g) reflected radiation as well as (d, h) canopy reflectance for multispectral channels 1 and 2. Results from the grassland site are shown in the left hand column, those from the forest site in the right hand column. Data that were selected (sel.) after quality filtering are shown as black filled circles, while data that was not selected (nsel.) are shown as grey filled triangles. An exception to this are panels (d) and (h) where for channel 1 the selected data are shown in red in order to visually distinguish them from the channel 4 data. In panel (c), the white reference (WR) panel observations averaged over all four resampled hyperspectral channels are shown as purple crosses. The standard deviations of the averaged ratios in panels (a)-(c) and (e)-(g) are not shown for the sake of clarity but were on the order of 10% or smaller of the mean values when averaged over the whole time series.

was, however, mostly observed in 2015, while the 2016 results were characterized by a much higher degree of stability. The downwelling ratio increased gradually in 2015, while the upwelling ratio decreased. This led to a stronger decrease from values of about 1.5 to 1 in the reflectance ratio due to the combined effects of up- and downwelling patterns. It appears that the upwelling ratio shows a somewhat parallel behavior as near-infrared reflectance in 2016

(Fig. 5.3f), which is also visible in the reflectance ratio (Fig. 5.3g). This variation was much smaller as the changes observed in 2015, however.

5.3.3 Continuous stability correction of hyperspectral observations

Grassland site The hyperspectral observations after the one-time calibration were partly lower than the multispectral ones but partly largely overestimated for all channels (Fig. 5.4a-d). The daily average ratios of multi- over hyperspectral observations averaged over all four channels were then used as continuous correction factors. This led to considerable improvements in the agreement between multi- and hyperspectral observations for all channels (Fig. 5.4a-d, Table 5.2).

The correction consistently reduced the considerable reflectance bias due to an overall tendency of overestimation to absolute values smaller or equal than 0.005 and 0.022 for channel 4 (Table 5.2). Squared Pearson correlation, r^2 , for reflectance improved from values smaller than 0.7 to values above 0.9 for all channels ($r^2 = 0.98$ for channels 1, 2 and 3). Root-mean-square error (RMSE) was considerably reduced to values smaller than 0.01 for channels 1 to 3 and 0.04 for channel 4 due to the combined effects of improved bias and higher correlation. The RMSE relative to the mean value (rRMSE) was smaller or equal to 14% after the continuous correction, while it was on the order of 50% before (Table 5.2).

The agreement between multi- and hyperspectral up- and downwelling irradiance measurements also showed clear improvements after the continuous correction (Table 5.2, Figs. 5.9, 5.10). rRMSE improved from 0.27 to 0.06 for upwelling and from 0.33 to 0.04 for downwelling observations when averaged over all four spectral channels. Improvements in r^2 values, however, were less clear as uncorrected observations already had very high values (0.98 for up- and 0.88 for downwelling uncorrected obs.; >0.995 for corrected obs.). rRMSE results indicated considerable differences between channels (detailed results not shown, see standard deviations in Table 5.2).

Forest site A general trend towards overestimation of hyper- compared to multispectral reflectance was observed at the forest site (Fig. 5.4e-h). Furthermore, channel 2 appeared to show considerably larger reflectance scatter than the other channels, which can also be

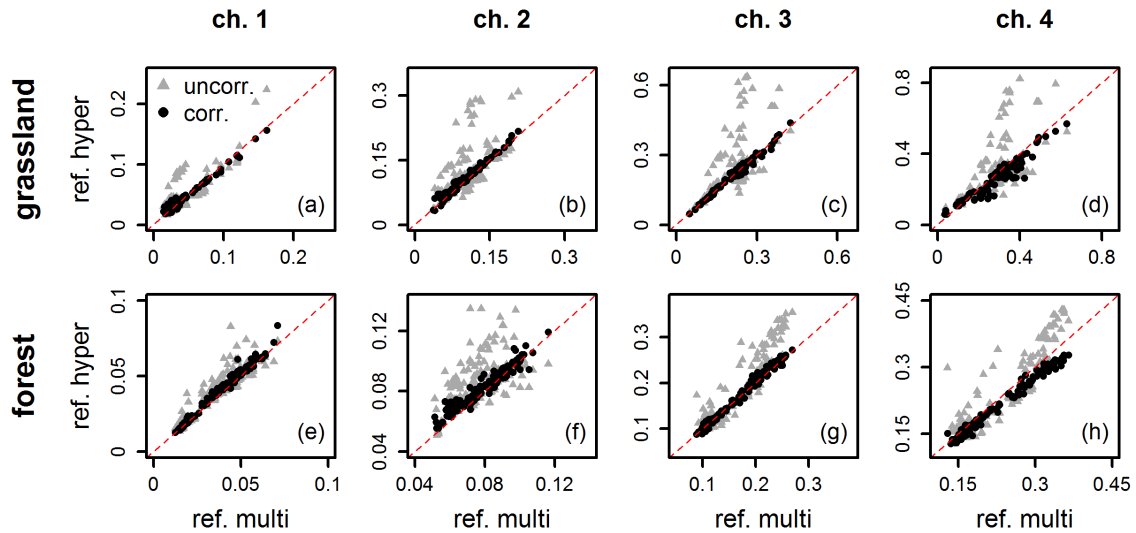


Figure 5.4: Daily averaged resampled hyperspectral reflectance ('ref.') vs. the corresponding multi-spectral reflectance for channels 1-4 and the grassland site in panels (a)-(d) and the forest site in panels (e)-(g). Resampled hyperspectral observations are shown both before (grey filled triangles) and after (black filled circles) continuous correction with the average over all four multispectral channels. The one-to-one line is shown as dashed red line.

seen in the low r^2 value of 0.3 (Table 5.2). The continuous correction led to considerable improvements in reflectance bias for channels 1 and 2, while bias increased almost two-fold for channel 4. Squared Pearson correlation increased for all channels, however, with values as high as 0.97 and 0.98 for channels 1 and 4 and 0.83 for channel 2. rRMSE could be reduced from about 20% before the continuous correction to less than 10% after the correction.

Up- and downwelling observations had very high r^2 values already before the correction (≥ 0.98) similar to the grassland site (Figs. 5.9, 5.10) but an increase was still observed for the correction ($r^2 \geq 0.996$). The improvement in terms of rRMSE was more notable with values improving by a factor of two to three to reach 0.11 for the average over all four upwelling channels and 0.06 for the downwelling ones (Table 5.2).

Site	Channel	r^2		RMSE		rRMSE		Bias	
		corr.	uncorr.	corr.	uncorr.	corr.	uncorr.	corr.	uncorr.
grassland	1 ref	0.98	0.69	0.005	0.020	0.12	0.47	0.001	0.008
	2 ref	0.98	0.34	0.008	0.056	0.07	0.54	0.005	0.026
	3 ref	0.98	0.39	0.010	0.114	0.05	0.53	0.001	0.044
	4 ref	0.90	0.43	0.041	0.129	0.14	0.45	-0.022	0.031
	1-4 up avg	1.00	0.98	-	-	0.06	0.27	-	-
	1-4 up sd	(0.001)	(0.002)	-	-	(0.004)	(0.08)	-	-
	1-4 dn avg	1.00	0.88	-	-	0.04	0.33	-	-
	1-4 dn sd	(0.001)	(0.006)	-	-	(0.02)	(0.006)	-	-
forest	1 ref	0.99	0.86	0.003	0.007	0.08	0.21	-0.002	0.004
	2 ref	0.93	0.30	0.005	0.018	0.06	0.24	0.003	0.010
	3 ref	0.99	0.88	0.006	0.038	0.04	0.25	0.001	0.18
	4 ref	0.99	0.85	0.017	0.037	0.08	0.17	-0.013	0.008
	1-4 up avg	1.00	0.99	-	-	0.11	0.19	-	-
	1-4 up sd	(0.001)	(0.007)	-	-	(0.05)	(0.08)	-	-
	1-4 dn avg	1.00	0.98	-	-	(0.06)	(0.13)	-	-
	1-4 dn sd	(0.001)	(0.003)	-	-	(0.02)	(0.03)	-	-

Table 5.2: Comparison of daily averaged multi- and resampled hyperspectral reflectance ('ref'), upwelling ('up') and downwelling ('dn') irradiance with and without continuous correction. E.g. '2 ref' refers to the reflectance observations of channel 2. For reflectance, results are given for all channels separately, for up- and downwelling irradiance only as averages with standard deviations in parentheses. Furthermore, up- and downwelling results are not given as absolute RMSE and bias due to the difference in channel FWHM. Squared Pearson correlation (r^2), RMSE, RMSE relative to the mean value (rRMSE) and bias values are shown. For the continuous method, all hyperspectral observations of a given day were scaled with the average ratio of multispectral channel 3 over correspondingly resampled hyperspectral observations on that day. The values without continuous correction were calibrated only on one day. The best performing values for each evaluation metric and site-channel case are printed in bold.

5.3.4 Comparison of hyperspectral reflectance observations with calibrated field spectrometer and example spectra during phase of rapid phenological changes

Measurements with a calibrated, portable field spectrometer were conducted at both sites in order to validate the calibration of the automated hyperspectral systems' reflectance. The multispectral observations agreed very well with the portable spectrometer at both sites

(Fig. 5.5a,b).

The reflectance spectrum of the automated hyperspectral system at the grassland site showed good agreement with the portable spectrometer in the visible range and the red-edge region of steep increase in reflectance. There was some discrepancy in the near-infrared range, however, where the automated system showed lower values that were at the edge of the lower one standard deviation range of the portable observations. Also, a more pronounced dip in the water absorption region at around 950 nm was observed for the automated system. At the forest site, the agreement between portable and automated system was better than at the grassland site, in particular in the near-infrared region where the automated observations fell inside the range of standard deviation of the portable measurements. The difference at the water absorption feature at 950 nm was also observed, however.

Example spectra of the increase in reflectance during the green-up in 2015 are presented in Fig. 5.5c-d for both sites. At the grassland site, the green-up started already at the end of February and the spectra during winter time still showed the characteristic shape for vegetation canopies with strong absorption in the visible range and much higher, relatively flat reflectance in the near-infrared. The near-infrared reflectance then increased from about 0.1 to more than 0.45 until the beginning of June (Fig. 5.5c). The reflectance in the 550 nm region characterized by slightly lower absorption of chlorophyll increased in parallel with the near-infrared reflectance. At the forest site, the green-up started in late March (or early April) when the reflectance spectrum only showed a weak indication of steeper increase at the red-edge region (about 700 nm) and was relatively high in the visible range. Until mid-June the reflectance in the visible range decreased, while it increased in the near-infrared range and also tended to become more spectrally flat.

5.4 Discussion

5.4.1 Potential and likely causes of calibration instability of hyperspectral systems

Our results showed that multispectral reflectance observations were rather stable (Fig. 5.2a), while there was considerable instability in the automated hyperspectral observations at both sites (Figs. 5.2a, 5.3), which raises the question of the causes for this. We observed that

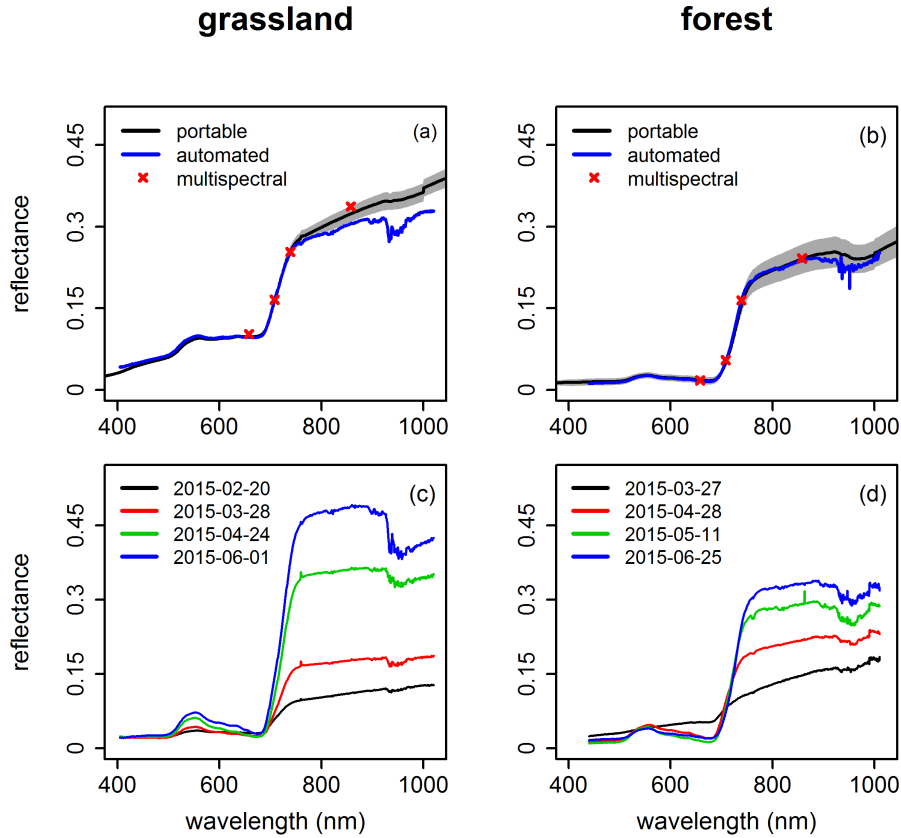


Figure 5.5: Validation of automated multi- and hyperspectral canopy reflectance observations with portable field spectrometer measurements (a, b) and examples of green-up in 2015 for (c) grassland and (d) broadleaf forest. The standard deviations from the mean of the portable observations are shown as grey shaded bands in panels (a) and (b), while standard deviations are not shown for the automated systems during the portable measurements due to the small values. Daily mean spectra are shown in panels (c) and (d).

the connections between different parts of the optical fiber path, e.g. between the shutters and the fibers or the fiber inside and outside the temperature control box, were very sensitive to mechanical disturbance such as tightening of the screwed connections. Furthermore, the bending of the optical fiber from the spectrometer to the tower mast arm can result in variable signal losses if the optical fiber is not completely fixed. This is expected to be a minor contribution at the grassland site as the mast arm was fixed. At the forest site, however, the mast arm was pivoted at about two-weekly time intervals in order to bring the installed instruments close enough to the tower for the purpose of maintenance.

As the instabilities strongly affected the reflectance ratio, i.e. affected the upward and down-

ward looking measurements in different ways, and the changes were consistent over the wavelength range, it can be excluded that the detectors of the spectrometer were the cause of it. suddenly occurring permanent signal loss in solar irradiance measurements (Fig. 5.3a) The suddenly occurring permanent signal loss in solar irradiance measurements at the grassland site (Fig. 5.3a) is expected to be caused by a fiber connection, potentially due to unintended mechanical impacts during maintenance activities at the site. The large drop in the upwelling ratio (Fig. 5.3b), in contrast, appears to have a different cause as the signal recovered after some time. Gradual changes are suspected to be mainly caused by variable losses due to fiber bending or small changes in fiber connections.

Furthermore, we observed that both up- and downwelling radiation fluxes contributed to the observed calibration instabilities in reflectance (Fig. 5.3). Instabilities and drift in the calibration of downwelling fluxes can be caused by partially damaged cosine correctors or impurities on their surface. This tends to decrease the signal but gradual increases of hyperspectral observations relative to the multispectral references were observed at both sites (Fig. 5.3a, e). Instabilities in the upwelling observations are less likely to be caused by impurities as the sensor faces downward. Differences between multi- and hyperspectral observations could then partly be caused by imperfect overlap of field-of-views (FOVs). There appear to be some indications for a small effect caused by this in our observations (Fig. 5.3b,f) and differences are expected to be larger at the forest site due to the off-nadir observation angle and larger distance between sensors and canopy. The main instabilities, however, did not show any relationship to seasonal vegetation patterns (Figs. 5.3d,h).

5.4.2 Spectral stability of reflectance correction factor

One important assumption for the presented correction approach is the spectral stability of the ratio of up- to downwelling observations. The probable causes for instabilities presented above (section 5.4.1), namely losses in fiber connections or errors in shutter timing, do not suggest a wavelength dependence. Fiber propagation or transmission losses are known to be wavelength-dependent (Paschotta, 2013) but the signal insertion losses at connections affects the fraction of the whole signal that enters the fiber and are therefore expected to be independent of wavelength. Furthermore, the high degree of consistency between the four resampled hyperspectral channels for each date of WR measurements (Fig. 5.2a) and the good performance of the continuous daily correction using the same correction factor for all

channels (Fig. 5.4, Table 5.2) are strong indications that the temporal instabilities were indeed predominantly limited to wavelength independent scaling of the whole spectra. Larger discrepancies after the continuous correction would be expected between the different re-sampled hyperspectral channels if the spectral shape of the correction factor was unstable. It is notable that the hyperspectral WR observations showed smaller CV between channels than the multispectral observations, which might be caused by small drift of the single detectors per multispectral channel while the resampled hyperspectral channels represent averages over a larger number of detector pixels (about 80 for channel 1 and about 20 for the other three channels). The small spectral differences in the correction factors between up- and downward measurements (Fig. 5.2b) are more likely caused by measurement errors due to e.g. not completely level WR panels, partial shading of diffuse contribution by structures on the tower mast arm, small changes in WR reflectance over time, especially in the visible range (Porcar-Castell et al., 2015) and by under- or overcorrection of the weak vegetation signal in the observations of the WR panel (section 5.2.3). However, changes in the optical properties of the diffusor of the hyperspectral systems over time due to exposure to rather extreme meteorological conditions or dirt on its surface could also change the transmission in a wavelength dependent way. The diffusors were therefore cleaned at regular intervals to reduce such effects.

5.4.3 Stability of multispectral measurements

Another important assumption of the presented correction method is the stability of the multispectral measurements. Firstly, there are theoretical reasons for their better stability based on the reasoning presented in section 5.4.1: no optical fibers are used for the type of sensor we used and the filters and detectors are typically mechanically fixed directly to the housing. Instability can then still result from a faulty detector and degradation of exposed surfaces through which the light enters the instrument but is unlikely caused by changes in the optical path between the point of light collection and the detector. Hyperspectral systems installed at eddy covariance towers, in contrast, typically use long optical fibers with potentially several connections between the point of measurement and the spectrometer, which are prone to signal losses. This is the main difference to the multispectral sensors when considering reflectance measurements as the other factors such as degradation of the detector or diffusor surface are relevant for both types of sensors.

Secondly, we also verified the validity of the assumption of a high degree of temporal sta-

bility of the multispectral sensors we used by applying three independent methods: First, a comparison of multispectral solar irradiance with broadband short wave downward radiation from pyranometers at the same sites (results not shown); second, a cross-comparison of multispectral sensors installed at the grassland and forest sites on selected cloud-free days (Fig. 5.8); third, WR measurements at the grassland site (Fig. 5.2a). The first method represents a continuous and the second method a quasi-continuous verification of solar irradiance stability, while the third included upwelling reflected radiation and hence was directly testing reflectance stability. All methods showed high temporal stability of the multispectral system (method 1: Dechant et al. (2018); method 2: Fig. 5.8; method 3: Fig. 5.2a). The problems with channels of the upward-looking multispectral sensors were limited to a single channel at each site and could be solved by replacing it with scaled observations of the neighboring channels (Dechant et al., 2018). This instrument issue is related to the case of faulty detectors as mentioned above and can, in principle, equally affect hyperspectral sensors. In the latter case, however, the issue tends to be more easily solved for both up- and downward looking observations due to the high number of partly redundant spectral channels. We used commercially available multispectral sensors and while the general reasoning of higher stability of multispectral sensors due to simpler design and operating principle is valid, products of other manufacturers or custom-built sensors built by researchers themselves might show different performances including lower stability.

5.4.4 Performance of the continuous correction of hyperspectral observations

The agreement between hyper- and multispectral reflectance observations after the continuous correction showed very good performance in terms of increase in correlation and bias as well as decrease in RMSE for reflectance as well as up- and downwelling radiation fluxes (Table 5.2, Figs. 5.4, 5.9, 5.10). In the case of channel 4 at the forest site, however, the bias was larger after the correction than before, despite enhancement of r^2 and RMSE. This appears to be caused by a slight but consistent bias of the corrected hyperspectral reflectance towards lower values than the multispectral reflectance (Fig. 5.4h) and could be caused by a small error of the spectral response calibration or the ratio of up- over downwelling radiation as observed with a WR panel (Fig. 5.1). There is no indication that this is caused by a temporal instability as the disparity could be effectively corrected by a temporally constant

scaling factor.

We found that the corrected reflectance spectra from the hyperspectral systems agreed well with the overall spectral pattern of the portable field spectrometer (Fig. 5.5a,b). However, a larger difference was observed for the near-infrared reflectance of the grass canopy (Fig. 5.5a). One potential explanation for the observed differences might be the presence of heterogeneities in the reflectance that could affect the multi- and hyperspectral sensors in a different way due to slightly different areas measured by both sensors despite a large overlap of field-of-view (FOV). This is more plausible for the difference between the portable and the automated hyperspectral systems as the former could not be measured at the same location due to practical reasons and a slight off-nadir angle was used to maximize overlap of the two FOVs. The higher errors of upwelling fluxes compared to the downwelling ones appear to further support this reasoning (Table 5.2, corrected results).

We pointed out in the methods section that the chosen time of the one-time calibration is somewhat arbitrary and we therefore evaluated its effect on rRMSE. We did this by comparing our reflectance results (Table 5.2) with those of the optimal one-time calibration factor equal to the average correction factor for the whole time series. We found improvement for the latter case as expected but it was limited to about 20% smaller RMSE at both sites compared to our chosen one-time calibration factor. This, however, is a much smaller effect than that of the continuous correction (RMSE reduction by factors larger than two and three for the forest and grassland site, respectively). These results therefore confirm that our selection of one-time calibration dates did not disproportionately inflate errors before continuous correction.

Another aspect is the influence of the discontinuity in the downwelling solar irradiance observations at the grassland site on the one-time calibration results. It is clear that it considerably increases RMSE and decreases r^2 compared to the forest site but could in principle be corrected in a discontinuous way. To quantify this effect and obtain more comparable results to the forest site, we tested the application of a discontinuity correction consisting of a constant scaling factor after the sudden drop (Figs. 5.4, 5.11). The results still showed considerable improvement for the continuous correction (Table 5.3, Fig. 5.12) and were indeed rather similar to the forest site in terms of RMSE and r^2 .

For practical purposes, it might be of interest to conduct a continuous calibration monitoring and correction with a multispectral instrument with fewer than four channels, potentially a single one. We therefore evaluated the performance of the continuous reflectance correction

using all single channels as well as various averages over pairs and groups of channels (Table 5.4). We found that even single channel corrections led to considerable improvements in r^2 for both sites but that multi-channel averages led to somewhat better results. There was no clear indication of worse performance of the reflectance results for continuous correction involving a downwelling multispectral channel that had been replaced with a neighboring one due to detector problems (section 5.2.3).

Our results highlight the importance of not only checking correlation measures for up- and downwelling fluxes as they tend to show very high correlation despite serious calibration issues (Table 5.2). Using relative RMSE measures for up- and downwelling fluxes as well as correlation measures for reflectance increases the sensitivity to calibration errors.

5.4.5 Comparison of different hyperspectral system set-ups with respect to reflectance calibration stability

The measurement set-up of hyperspectral systems affects the stability of reflectance observations. Thus, an important advantage of the WhiteRef hyperspectral system (Sakowska et al., 2015) is that it uses a single spectrometer and measures the WR before (and after) each target measurement similar to the mode of operation used with portable instruments in field campaigns. This ensures continuous correction between upward and downward looking measurements. A certain effort in installation and maintenance is required, however. The hyperspectral irradiometer system (Meroni et al., 2011) avoids some of the issues we observed in our DFOV system by periodically changing the FOV from upward to downward looking configurations in a mechanical way. However, the mechanical movement might also affect the reflectance calibration due to the effects on the fiber optic cable. The system by Leuning et al. (2006) reduces this potential source of error by having a fixed optical fiber and only moving mirrors and shutters. A further limitation with the hyperspectral irradiometer system is that it can only measure bihemispherical reflectance and not hemispherical-conical reflectance, which is commonly measured with field spectrometers. Using two different spectrometers to measure the upward and downward radiation as in the system by Huber et al. (2014) is not only more expensive than a system based on a single spectrometer system but also does not resolve the reflectance calibration issue of systems with a single spectrometer and fiber switches. Moreover, systems using two spectrometers face cross-calibration issues with regard to the wavelength scale.

5.4.6 Limitations of our study

One limitation of our study is that the temporal stability of repeated spectral response calibrations was not assessed directly by either repeating the cross-calibration with the portable field spectrometer or a calibration light source several times during the study period. Temporal changes of the spectrometer spectral response can therefore not be strictly excluded. The consistency with the multispectral observations, however, is a strong indication for stability as it appears very unlikely that the spectral response of the two independent measurement systems based on entirely different instrument types would show parallel changes over time.

Another limitation is that the higher stability of multispectral observations was only directly assessed in terms of reflectance but not in terms of absolute radiometric calibrations of up- and downwelling radiation fluxes separately. Large temporal changes of the latter could be excluded, however, by comparing with high quality pyranometer observations of shortwave downward radiation (results not shown) as well as the inter-comparison between the multispectral downwelling irradiance observations at the two sites (Fig. 5.8). In addition, a parallel temporal change in both up- and downward facing multispectral sensors appears unlikely as they are fully independent instrument units.

It should also be noted that the correction of the hyperspectral WR observations for a contribution of the vegetation signal was not applied to the multispectral observations although the latter might in principle be affected in a similar way. Only one detector element for each channel is used in the multispectral sensor, and hence one cannot distinguish with certainty if the spectral pattern of the WR observations is due to a vegetation signal or calibration bias. We did, however, observe consistently higher values of channel 3 compared to channel 2 for the up/down ratio of the WR observations (results not shown), which is an indication that a small vegetation signal was also present in the multispectral WR observations. This could explain the larger standard deviation for the multispectral observations for individual dates and therefore we also compared the uncorrected multispectral with the uncorrected hyperspectral results. We found that the uncorrected hyperspectral results showed clearly higher standard deviation than the corrected ones for individual dates but that the overall dominance of the variation between dates was still evident (Fig. 5.13). This implies that our approach of correcting the hyperspectral WR observations did not affect the main conclusions drawn from the WR observations.

5.4.7 Implications of our results and different correction methods

We observed considerable instabilities in both up- and downwelling hyperspectral observations that were at least partly nonlinear and made a continuous correction necessary. We believe, therefore, that especially custom-made systems that are assembled out of commercially available components by the research institutes themselves should be carefully monitored either with sufficiently frequent campaign-based calibration measurements or with partly redundant or complementary continuous measurements. The former might not always be possible due to practical limitations related to either the remoteness of the site or available manpower.

Furthermore, care needs to be taken to obtain calibration measurements of sufficiently high quality as we found some issues with the WR measurements (section 5.2.3). A calibration light source might be helpful for this purpose but other technical difficulties for this type of observation might arise such as instability of the light source. For the continuous approach using partly redundant or complementary observations, multispectral measurements using relatively low-cost sensors (Balzarolo et al., 2011; Jin and Eklundh, 2015; Porcar-Castell et al., 2015) can be used for this purpose for both upward and downward looking observations, as we showed in this study.

Given spectral stability of the hyperspectral system, we could demonstrate that even an instrument with a single multispectral channel and stable operation would be enough for continuous calibration monitoring and correction. This suggests using continuous measurements of incoming photosynthetically active radiation (PAR) or shortwave downward radiation (SWDR) measured with pyranometers in a range of approximately 400 to 3000 nm to monitor solar irradiance stability. These measurements are routinely conducted at eddy covariance sites and supposing the range of the hyperspectral system covers the range of PAR or SWDR instruments, a direct comparison is possible. Otherwise, large instabilities might still be detected, while an accurate correction might be more difficult. For the monitoring and correction of the upwelling solar radiation reflected from the vegetation canopy, however, using PAR sensors or SWDR pyrgeometers requires compatible FOVs. Downward looking PAR sensors and SWDR pyrgeometers typically have hemispheric FOVs, which implies that the direct comparison with narrow angle conical observations is not possible. The effects of different FOVs may be large, especially in case of considerable optical heterogeneity in the

footprint. Moreover, using broadband sensors such as PAR or SWDR does not provide the possibility to also monitor spectral stability as with a multispectral sensor.

Our approach could also be used to correct instabilities of hyperspectral systems on the diurnal scale if necessary, potentially by not applying daily average correction factors but instantaneous ones based on the closest multispectral observation to each hyperspectral one. This might be relevant for studies of sun-induced chlorophyll fluorescence (SIF) or photochemical reflectance index (PRI) where diurnal patterns are related to photosynthetic uptake of CO_2 and a daily average would not yield the required information.

5.5 Conclusions

We observed large instabilities in the radiometric and reflectance calibrations of two automated hyperspectral measurement systems of vegetation canopy reflectance over a period of about 16 months. This and the nonlinear behavior of the observed instabilities confirmed that it is, in general, necessary to continuously monitor the temporal stability of these types of unattended optical measurement and apply corrections if needed. We found that multispectral observations at the same sites were more stable and could be used for a continuous calibration correction that led to large improvements such that relative errors between multispectral and resampled hyperspectral channels on the order of 1% at wavelengths in the visible and near-infrared ranges could be achieved. Together with results from white reference measurements, this confirmed that the temporal instabilities were predominantly limited to a spectrally constant scaling of the spectra rather than changes in the spectral shape.

Campaign-based calibration measurements are helpful to detect certain problems, but are in general not enough to either detect all problems or continuously correct the observations if necessary. Particularly for dual-field-of-view type hyperspectral measurement systems based on a fiber switch, it is necessary to closely monitor the correction factor between the upward and downward looking observations and correct for changes on a continuous basis, if possible. We suggest the use of stable multispectral sensors for continuous calibration monitoring and correction but complementary measurements that are conducted routinely at eddy covariance sites can also be used in principle, especially for the downwelling solar irradiance. In the absence of continuous calibration monitoring, increased frequency of calibration measurements could, of course, alleviate some of the potential problems. Our findings can contribute to improve comparability of automated hyperspectral observations between sites and

help obtain well calibrated data for model applications and satellite product validation.

Supplemental Material

Site	Channel	r^2		RMSE		rRMSE		Bias	
		corr.	uncorr.	corr.	uncorr.	corr.	uncorr.	corr.	uncorr.
grassland	1	0.98	0.84	0.005	0.013	0.12	0.32	-0.001	0.003
	2	0.98	0.57	0.008	0.024	0.07	0.24	-0.005	0.004
	3	0.98	0.57	0.010	0.054	0.05	0.25	-0.001	0.019
	4	0.90	0.62	0.041	0.081	0.14	0.28	0.022	0.046

Table 5.3: Results analogous to Fig. 5.4 for the grassland site but here the ‘uncorr.’ case indicating no continuous correction was corrected for the discontinuity of irradiance as shown in Fig. 5.11. The best performing values for each evaluation metric and site-channel case are printed in bold.

Site	Correction Channel	r^2 of evaluation channel			
		1	2	3	4
grassland	1	-	0.93	0.8	0.64
	2	0.98	-	0.96	0.82
	3	0.96	0.95	-	0.91
	4	0.88	0.77	0.90	-
	1-4	0.98	0.98	0.98	0.90
	1-3	0.99	0.99	0.96	0.81
	1-2	1.00	0.98	0.90	0.74
	3-4	0.93	0.88	0.97	0.98
	1	-	0.85	0.97	0.96
	2	0.98	-	0.95	0.94
forest	3	0.97	0.83	-	0.98
	4	0.96	0.79	0.98	-
	1-4	0.99	0.93	0.99	0.99
	1-3	0.99	0.93	0.99	0.99
	1-2	0.99	0.97	0.97	0.96
	3-4	0.97	0.84	0.99	1.00

Table 5.4: Comparison of continuous correction results for reflectance based on the correction factor of different multispectral channels. The best performing values for each evaluation channel and single-channel or multi-channel correction (average over several channels, e.g. 1-3) case are printed in bold for each site. Results are comparable to Table 5.2.

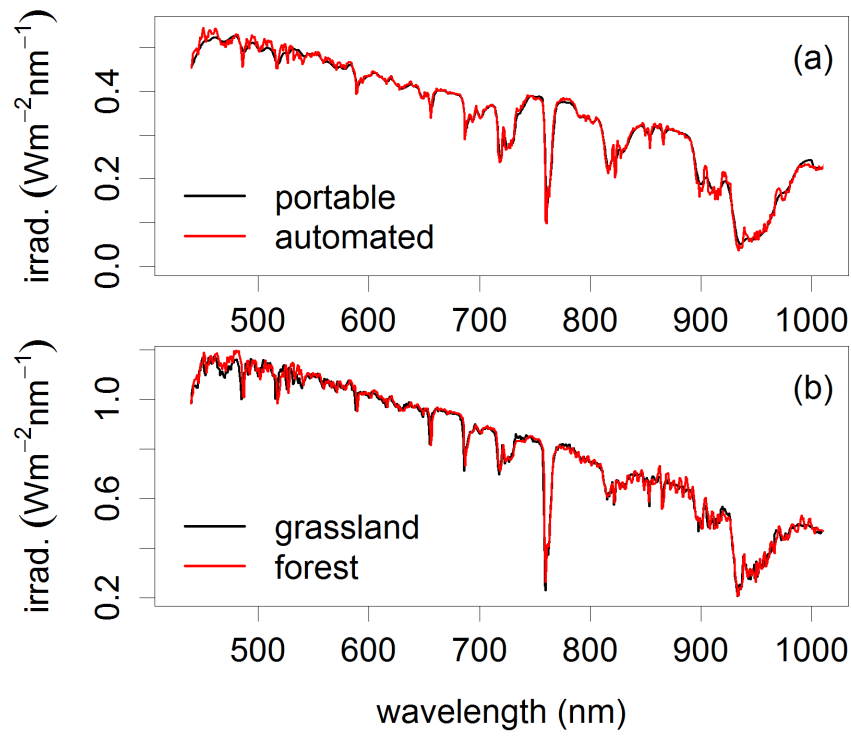


Figure 5.6: Reference hyperspectral solar irradiance measurements compared to calibration results for (a) the grassland site and (b) the forest site. At the grassland site a portable field spectrometer was used to calibrate the automated system while at the forest site the observation of the calibrated grassland system was used on a clear-sky day. Time and date differed for the two calibrations, hence the irradiance spectra for (a) and (b) differ in magnitude and shape.

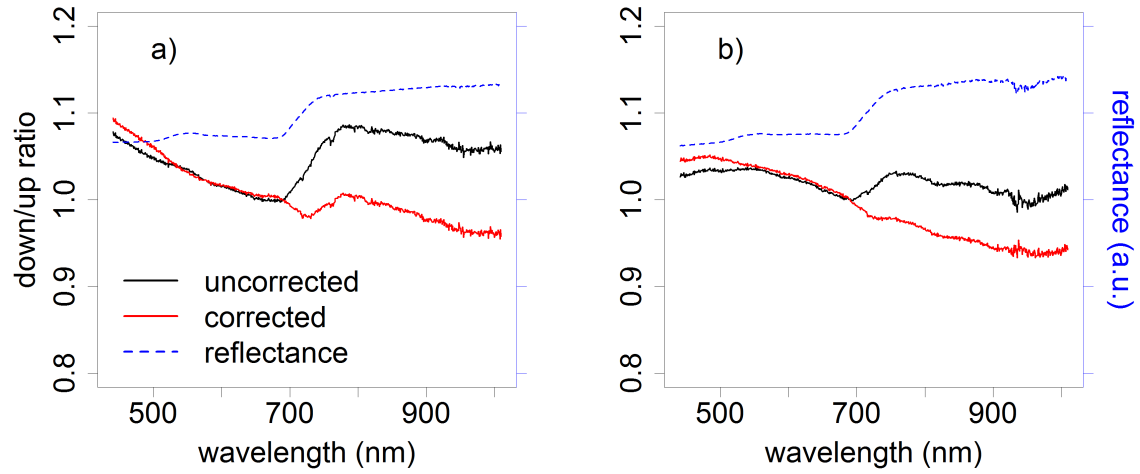


Figure 5.7: Comparison of uncorrected white reference (WR) panel measurement of the ratio of downward looking over upward looking observations for the same signal, the measurement corrected by subtracting a fraction of the vegetation upwelling radiance from the upwelling WR signal and a reflectance spectrum of the same day in arbitrary units (a.u.) as visual reference of the red-edge signature. In a) the measurement at the grassland site in March 2015 the red-edge feature could not be completely corrected, while in b) the measurement in July 2016 shows less artefacts after correction. Both corrected and uncorrected ratios are normalised at 690 nm for easier visual comparison.

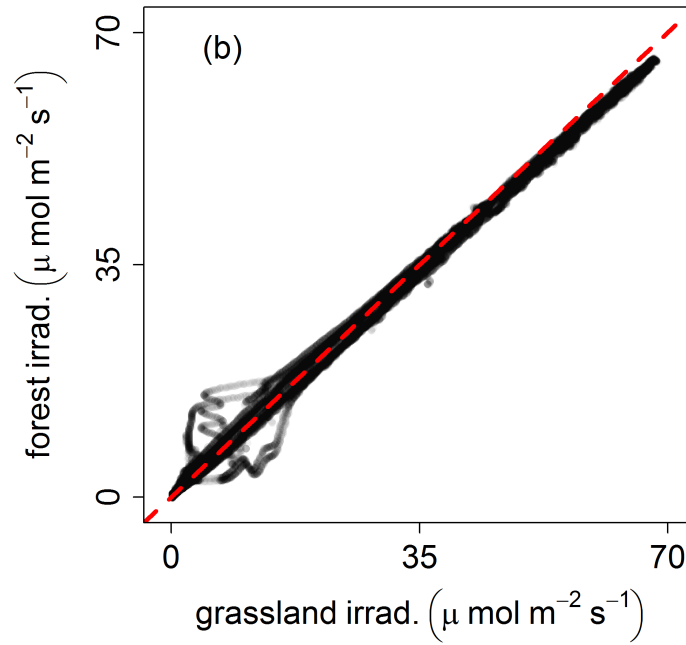


Figure 5.8: Comparison of optical observations of the grassland site and the forest site on seven sunny days in 2015 and 2016. The comparison is shown for multispectral observations of channel 2 (center wavelength at 708 nm). The pattern for low irradiance is related to partial sensor shading by other instruments. The red dashed line is the one-to-one line. The observations are shown as partly transparent filled circles.

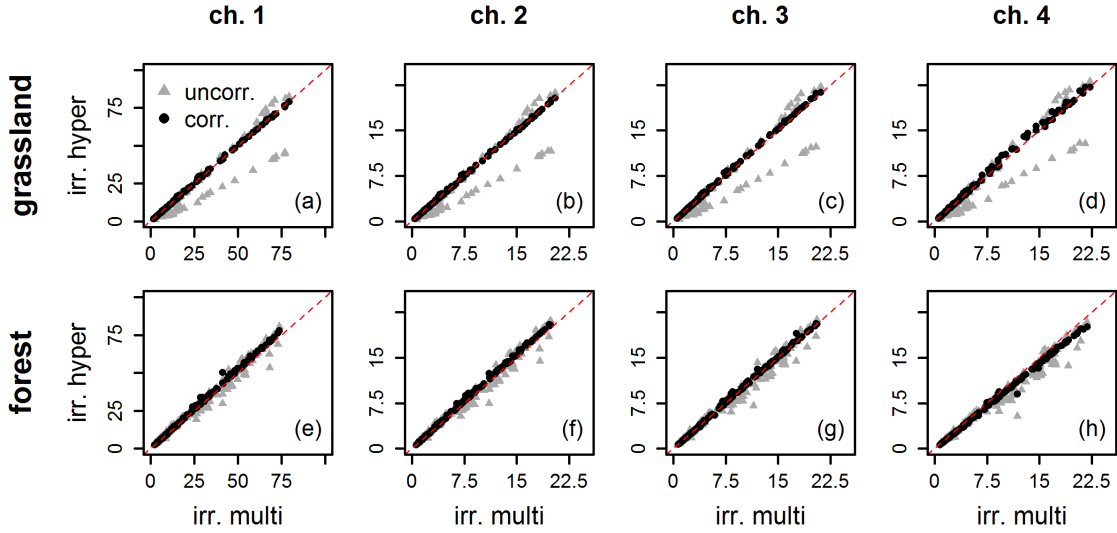


Figure 5.9: Results analogous to Fig. 5.4 but for downwelling solar irradiance in units of $\mu\text{mol m}^{-2} \text{s}^{-1}$ instead of reflectance. Values with only the one-time calibration are shown as grey triangles and values after continuous correction using multispectral observations as black circles. The one-to-one line is shown as dashed red line.

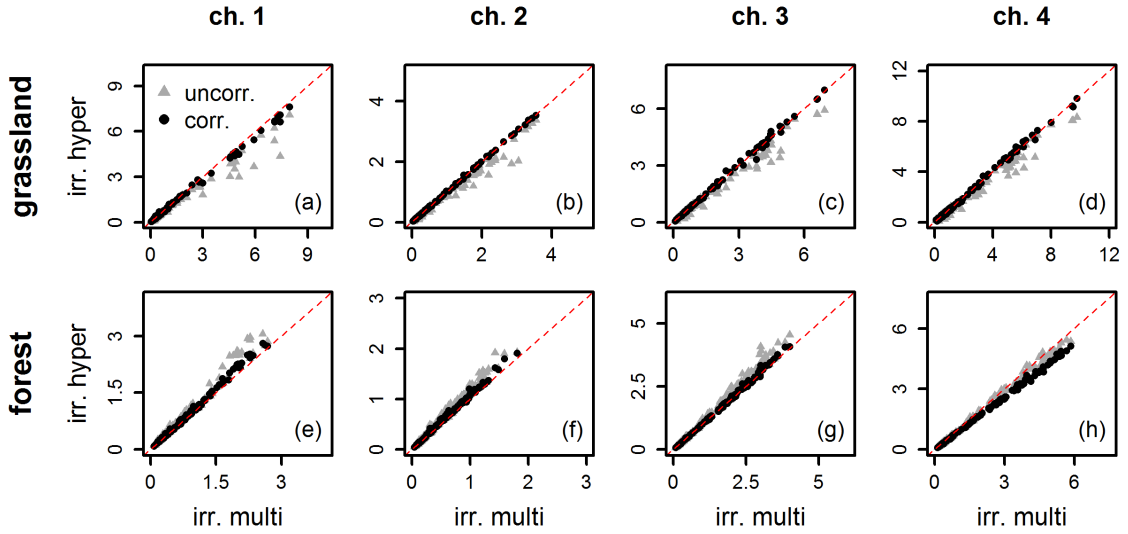


Figure 5.10: Results analogous to Fig. 5.4 but for upwelling reflected solar irradiance in units of $\mu\text{mol m}^{-2} \text{s}^{-1}$ instead of reflectance. For better comparison with the downwelling case, the observed upwelling radiance was multiplied by π to obtain units of irradiance. Values with only the one-time calibration are shown as grey triangles and values after continuous correction using multispectral observations as black circles. The one-to-one line is shown as dashed red line.

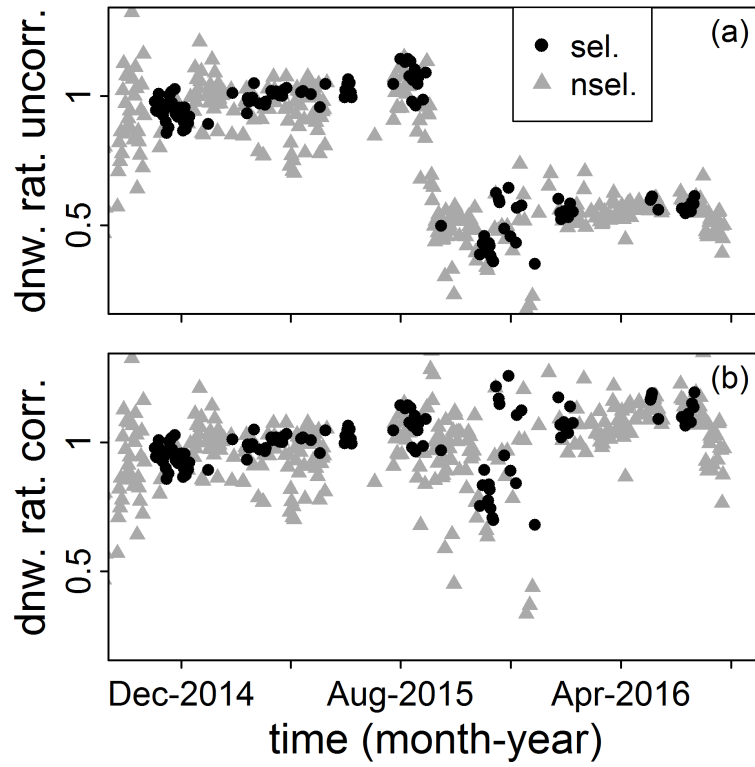


Figure 5.11: Comparison of (b) discontinuity corrected and (a) uncorrected ratio of hyperspectral over multispectral downwelling solar irradiance observations at the grassland site. For the discontinuity correction, a simple scaling factor was applied to attempt to correct the discontinuity in the data. Data that was selected (sel.) after quality filtering are shown as black filled circles, while data that was not selected (n.sel.) are shown as grey filled triangles.

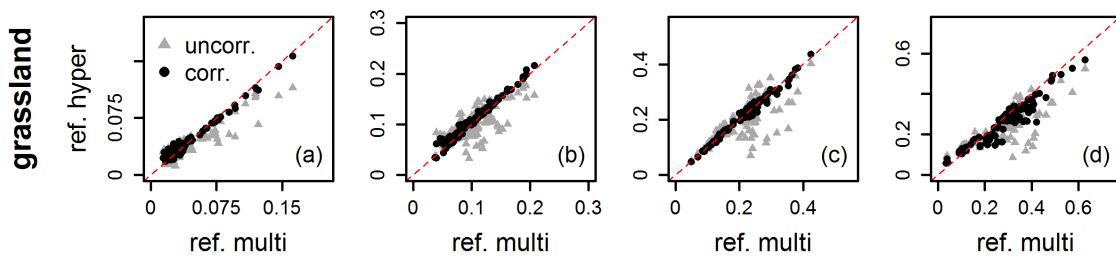


Figure 5.12: Results analogous to Fig. 5.4 for the grassland site but with correction of the discontinuity of irradiance as shown in Fig. 5.11 and Table 5.3. Panels (a) to (d) show the results for multispectral and hyperspectral resampled channels 1 to 4.

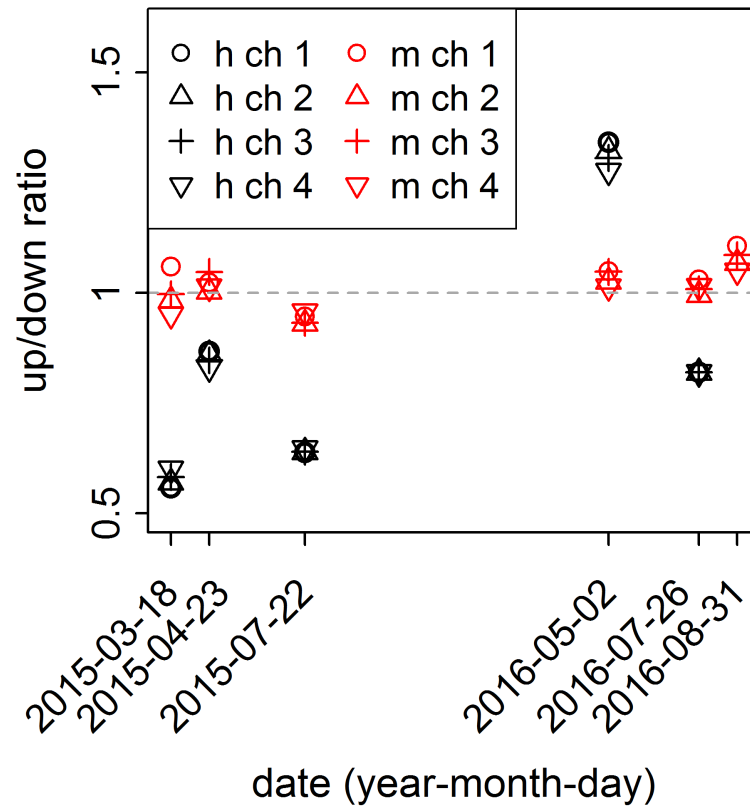


Figure 5.13: Results of the white reference observations analogous to Fig. 5.2(a) but without the correction of the vegetation signal in the hyperspectral observations.

6 Discussion

Altered phenological behaviour has been detected in ground observational (Füssel et al., 2012; Intergovernmental Panel on Climate Change, 2014; Fu et al., 2014) and remote sensing data (Myneni et al., 1997; Barichivich et al., 2013; Fu et al., 2014). Subsequent consequences for plants and ecosystems are numerous (see Chapter 1) and the future impact of these phenological changes is still a matter of debate. The first part of the discussion assesses whether the assumed climate warming is depicted by phenological models in correspondingly earlier bud burst (BB) dates. Extensive phenological and meteorological information is used to answer this research question. Consequently, monitoring phenology, including human observations and remote sensing products, is crucial in order to analyse the present state of vegetation and improve our understanding of the complicated mechanistic interactions between chilling requirements, warming spring temperatures and photo-period. Satellite-based phenological products could be beneficial for large scale climate change analysis, but the large number of methods available to acquire phenology products from satellite observations hampers inter-comparability between studies. The second part of the discussion deals with the applicability of methods extracting phenological metrics from satellite data on different land-use classes. The development and improvement of such methods relies on the knowledge about factors influencing the satellite signal and subsequently extracted phenological metrics. Optical measurements at ground level may bridge the gap between classic phenological observations of individual plants and satellite-based measurements. The third part of this discussion deals with the establishment of a satellite product validation site and their benefits for satellite phenology depiction. The operation of such a site includes the maintenance of the optical sensor systems to ensure data quality. One of the main challenges in this context is sensor calibration. The fourth part of the discussion examines the benefits of a continuous

calibration compared to the typically used campaign-based approach. Thus, this discussion covers the path from observation to area-wide prediction of phenology, including calibration of optical sensor systems and validation of plant phenology products.

6.1 Modelling bud burst in a future warming climate

The Intergovernmental Panel on Climate Change (2014) expects an earlier onset of spring phenological events in humid temperate and boreal forests, with some species showing a reduced response to global warming due to photo-period and chilling requirements. The key findings of the study presented in Chapter 2 outline a BB advancement of tree species as long as temperatures will not exceed a species specific temperature sum limit and photo-period as well as chilling requirements are fulfilled. A mean BB trend of -0.05 up to -0.11 days per year was detected, depending on species. Nevertheless, results show a stagnation of BB advancement in a warming climate due to limited light level and unfulfilled chilling requirements. These results are in line with other recent studies on the phenological response of tree species to global warming, reporting a decreased rate of BB advancement (Morin et al., 2009) and a reduced sensitivity of spring phenology to temperature (Fu et al., 2015b,a). In addition, previous modelling studies found a considerable higher trend towards earlier BB for the 1950s to the 2000s (Menzel and Fabian, 1999; Chmielewski and Rötzer, 2001; Menzel, 2013b).

There is evidence that the phenological response to climate warming is attributable to species dependent temperature (Morin et al., 2009; Fu et al., 2015b), photo-period (Caffarra et al., 2011; Basler and Körner, 2012) and chilling requirements (Laube et al., 2014; Lange et al., 2016). Further, differences between annual and perennial as well as between wild and agricultural plants have been reported (Estrella et al., 2009). Many inter-specific interactions in ecosystems depend on the synchrony of phenological events (Menzel, 2013b), e.g herbivores and host plants, timing of predators and prey peak abundance, as well as flowering and pollinator abundance (Intergovernmental Panel on Climate Change, 2014; Park, 2016). Consequently, phenological shifts might affect the competitive balance between species (Vitasse et al., 2011; Lange et al., 2016) and potentially impact agriculture and horticulture (Chmielewski, 2013). There is evidence that climate change may lead to an increased risk of frost damage (Kramer et al., 1996; Füssel et al., 2012; Chuine et al., 2013) and an enhanced

occurrence of pests (de Beurs and Henebry, 2013; Chmielewski, 2013). Subsequently, plant survival and productivity is decreased (Chuine et al., 2013). An enhanced knowledge about phenological modelling and thus improved phenology predictions might support the adaptation of strategies in managed ecosystems, e.g. by frost and insect damage risk assessments (Kramer et al., 1996; Chuine et al., 2013; de Beurs and Henebry, 2013), yield estimation of crops, the choice of varieties in agriculture, crop rotation management and the estimation of possibilities for catch cropping (Chmielewski, 2013; Chuine et al., 2013; Menzel, 2013b). Large scale modelling of tree phenology enables the evaluation of recent and future annual vegetation growth in forest ecosystems and consequently altered carbon budgets (Fu et al., 2014). Subsequently, it potentially provides spatially inclusive and comprehensive validation data for satellite phenology estimations in ecosystems where ground truth data is only sparsely available.

The stagnation of BB advancement has also been reported by remote sensing studies using satellite data (Fu et al., 2014). However, respective trend magnitudes varied considerably (White et al., 2009; Fu et al., 2014). Management and land-use changes as adaptation strategy to climate change (Chmielewski, 2013), the heterogeneity of vegetation within pixels (Fu et al., 2014), the choice of algorithms and their disparate applicability for land-use classes (see Chapter 3) as well as vegetation disturbances might explain the sometimes weak correlations between modelled plant phenology products and satellite estimates (White et al., 2009; Menzel, 2013b).

6.2 Comparison of methods extracting phenological metrics from satellite data

Numerous methods extracting phenological metrics from satellite data are available. Recent inter-comparison studies found significant variations in spring phenology estimated by different methods (White et al., 2009; Cong et al., 2012; Fu et al., 2014). The results presented in Chapter 3 suggest a disparate applicability of different methods across landscape compositions and are thus in line with these studies. However, average start-of-spring trend estimates are consistent across all methods. Averaged trend magnitudes (0.6 to -1.1 days/year) are in line with other studies depicting spring phenology trends using one methodology and averaging all land-cover types, such as Zhou et al. (2001) with -0.36 days/year between 1982 and 1999, Stöckli and Vidale (2004) with -0.54 days/year between 1982 and 2001 and Fu

et al. (2014) with -0.45 days/year between 1982 and 2011. Studies analysing the period after 2000 with the AVHRR dataset GIMMS found a delayed or reversed trend (Jeong et al., 2011; Piao et al., 2011; Fu et al., 2014). This may be related to species specific trend behaviour (Fu et al., 2014) in combination with the broad spatial resolution (8 km) of GIMMS, consequently integrating different land-use classes with disparate trend behaviour, temporally changing surface elements, land-use practices or managements within a pixel (Reed, 2006; Doktor et al., 2009). Further, recent studies suggested that the quality of the GIMMS dataset may have induced the reversed trend (Zhang et al., 2013; Fu et al., 2014). The MEDOKADS dataset used in Chapter 3 does not suffer from these quality issues.

In addition, the extracted spring phenology metrics for forested areas are validated with ground observations from an extensive phenological network and model results for tree species BB. Satellite trend estimates generally agree with the ground data and model results. Although trend estimates are consistent across methods, differences across land-use classes and mixtures thereof are still evident. The tested methods show a mean trend differences across land-use classes of up to one day per year. The interquartile range is generally lowest for broad-leaved forest and highest for agro-forest and annual crops. Further, the theoretical framework used in Chapter 3 outlines the sensitivity of trend estimates to the number and proportion of land-use classes within a satellite sensor pixel. Consequently, methods extracting spring phenology from satellite data vary in performance depending on landscape composition and thus estimate varying trends.

Ground observations and modelling are highly valuable for the analysis of spring phenology change due to extensive phenological records (Fu et al., 2014). Satellite estimates, on the other hand, offer a large spatial coverage, but describe the seasonality of reflectance characteristics of vegetated landscapes rather than particular species (Henebry and de Beurs, 2013; Fu et al., 2014). Thus, large discrepancies are encountered when comparing both approaches (Richardson et al., 2013; Fu et al., 2014). Consequently, linking in-situ observations and satellite products on each land-use class is crucial to understand the relationships between optical information and ecological, as well as physiological processes (Lange et al., 2017) and to address the uncertainty related to different sensors in phenological studies (Badeck et al., 2004; Fu et al., 2014).

6.3 Validation of phenology products with near-surface sensor-derived data

New satellite generations, such as Copernicus Sentinel-2, may overcome the issue of integrating different land-use classes within one satellite pixel by their fine spatial resolution (Lange et al., 2017). Nevertheless, non-vegetational effects, such as atmospheric conditions, snow cover, viewing geometry, illumination conditions and signal distortions, influence remote sensing measurements (Soudani et al., 2008; Eklundh et al., 2011). The results presented in Chapter 4 confirm these effects and outline their disparate influence on phenological metrics extraction. Generally, in-situ and satellite-based phenological metrics agree with each other. Sentinel-2 based phenological metrics show a slightly better agreement with in-situ measurements than MODIS based metrics. Further, temporal mismatches between different sensor types may influence the consistency of phenological metrics (Soudani et al., 2008; Lange et al., 2017). These results assume a benefit of using satellite data with finer spatial and temporal resolution and are thus in line with other recent remote sensing studies (Fisher and Mustard, 2007; Soudani et al., 2008; Munyati and Mboweni, 2013).

Despite the progress in satellite data resolution and quality, the influence of non-vegetational effects could not be fully removed even with modern correction procedures (Lange et al., 2017). Thus, dynamic filtering algorithms like BISE are still beneficial to generate reliable phenological time series. Ground measurements are highly valuable to examine these influencing factors, which occur mainly during periods with low signal-to-noise ratio introduced by bad illumination conditions and cloud or snow cover. Near-surface remote sensing data may help to interpret the observed patterns biologically, since the sensors field of view contains a well-defined group of organisms in contrast to a mixture of species in satellite pixels (Soudani et al., 2012; Richardson et al., 2013). Optical measurements are not influenced by human observers like traditional phenology observations (Eklundh et al., 2011; Richardson et al., 2013). Thus, they may bridge the scale between classic phenology observation of individual plants and satellite-based remote sensing of land surface phenology (Richardson et al., 2013; Lange et al., 2017).

The set-up of such a near-surface remote sensing site mainly depends on purpose and budget. The study presented in Chapter 4 uses a sensor system measuring a hemispherical-conical reflectance factor (HCRF, Nicodemus et al. (1977); Porcar-Castell et al. (2015)) with one hyperspectral sensor in DFOV mode and two multispectral sensors in DFOV mode (Porcar-

Castell et al., 2015). Three instrument configurations are commonly used in in-situ field measurements: hemispherical-conical, bi-conical and bi-hemispherical (Porcar-Castell et al., 2015). The hemispherical-conical set-up is chosen in order to measure a well-defined area of the canopy while including down-welling irradiance from different solar zenith angles and diffuse illumination (Porcar-Castell et al., 2015). The bi-conical set-up is another viable option for this purpose, although it relies on a diffuse white reference panel to provide the reference measurement (Porcar-Castell et al., 2015; Sakowska et al., 2015). This panel has to be kept clean, which can be a challenge in outdoor conditions. The bi-hemispherical set-up includes up-welling radiance from all viewing directions and samples a wider area, consequently being more sensitive to variations in illumination geometry (Meroni et al., 2011; Porcar-Castell et al., 2015).

The study presented in Chapter 4 chose a DFOV set-up in order to measure up-welling and down-welling radiation near-simultaneously without the need for moving parts, which are prone to degradation (Meroni et al., 2011; Sakowska et al., 2015). This set-up usually involves two sensors and thus requires sensor inter-calibration (Porcar-Castell et al., 2015). New systems, e.g. the Piccolo Doppio (MacArthur et al., 2014), use bifurcated cables in order to install one spectrometer in DFOV mode and thus save the costs of an additional spectrometer and avoid sensor inter-calibration and moving parts. The study presented in Chapter 4 uses the latter set-up for the hyperspectral sensor. However, two multispectral sensors in DFOV mode are installed due to their low costs and calibration efforts. Whereas the installation of multispectral sensors is sufficient for phenological pattern analysis and the comparison with satellite channels with respective configuration, the information acquired from the hyperspectral sensor provides some benefits: The high spectral resolution allows for the validation of other remote sensing products (Porcar-Castell et al., 2015), such as chlorophyll content and leaf area index (Doktor et al., 2014), as well as the possibility to model, for example, plant productivity (Gamon, 2015; Hilker et al., 2008a). Further, the capability to generate multispectral information with different central wavelength and FWHM allows for the validation of different optical satellite missions (Lange et al., 2017). Upcoming satellite missions with different band configurations consequently benefit from satellite validation sites equipped with hyperspectral sensors.

These numerous possible set-ups and configurations, including the extent of the optical footprint, viewing geometry, spectral configuration, sampling intervals and calibration status, demand for a standardisation of optical measurements at satellite validation and EC sites (Balzarolo et al., 2011). Making reproducible measurements requires a detailed assessment

and documentation of data quality (Balzarolo et al., 2011). Further, characterisation of instruments is crucial since instrument issues, including linearity of spectrometer response, cosine response of diffusers, signal-to-noise ratios as well as radiometric and spectral calibration and their stability, may influence measurements (Eklundh et al., 2011; Porcar-Castell et al., 2015). Although the characterisation should be done on a regular basis, usually at least annually, the procedure can be resource and time consuming (Pacheco-Labrador and Martín, 2015; Porcar-Castell et al., 2015). In-situ characterisation, calibration and validation methods, avoiding the dismantling of automated systems, are thus recommended to ensure temporal stability of measurements and calibration (Pacheco-Labrador and Martín, 2015).

6.4 Continuous calibration of unattended near-surface sensor-systems

Optical sensor systems typically require two types of calibration: radiometric and spectral calibration (Balzarolo et al., 2011; Porcar-Castell et al., 2015). These calibrations are not stable over time, thus calibration measurements have to be repeated on a regular basis (Eklundh et al., 2011). However, the optimal calibration frequency is still a matter of debate (Balzarolo et al., 2011; Porcar-Castell et al., 2015). The study presented in Chapter 5 clearly shows the need for continuous calibration monitoring of unattended multi- and hyperspectral measurements, since campaign-based approaches can be not enough to either detect certain problems or continuously correct the observations if it is necessary. The results outline that a continuous calibration can be achieved at (relatively) low cost by utilising multispectral or pyranometer measurements, which are usually conducted routinely at EC sites where unattended measurement systems of vegetation canopy optical properties are typically located.

The observed instabilities in radiometric and reflectance calibrations mainly affect applications that require absolute reflectance calibrations, especially radiative transfer model inversion and the validation of satellite reflectance measurements, and studies relying on absolute calibration on both upward and downward looking channels, such as studies of sun-induced fluorescence. Further, instabilities in the spectral calibration or wavelength-dependent instabilities in radiometric and reflectance calibrations, although not observed in the study presented in Chapter 5, affect applications utilizing ratio or normalized differences indices, since they rely on information acquired in specific wavelength regions (Eklundh et al., 2011; Porcar-Castell et al., 2015). Thus, the calibration procedure presented in Chapter 5 can con-

tribute to improve data quality of optical measurements and is therefore crucial for studies of phenology.

Reducing the maintenance effort necessary for the operation of unattended optical measurement systems is a timely challenge given the increasing demand for optical measurements at EC and satellite product validation sites (Balzarolo et al., 2011; Eklundh et al., 2011; Pacheco-Labrador and Martín, 2015). Maintenance of such systems typically includes cleaning of fore-optics and reference panels (Porcar-Castell et al., 2015), in-situ and laboratory calibrations (Eklundh et al., 2011), as well as dealing with problems encountered in outdoor conditions, such as degradation through weather conditions and animal damage. The automation of calibration stability monitoring and error detection is thus highly beneficial and helps obtaining reproducible measurements. Regular reporting of the obtained information may improve comparability to measurements acquired at other sites (Balzarolo et al., 2011).

6.5 Synopsis of modelling, monitoring and validation of plant phenology products

The synoptic assessment of plant phenological changes in a temperate climate presented in this thesis reveals interactions of climate change and vegetation dynamics. The use of a physiology-based phenological model fitted to the extensive phenological database of the German Weather Service and driven by several climate scenarios enabled the forecasting of deciduous forest tree BB. Statistics on historical data used in Chapter 2, other recent ground observational (Füssel et al., 2012; Intergovernmental Panel on Climate Change, 2014), modelling (Vitasse et al., 2011; Fu et al., 2012; Migliavacca et al., 2012) and remote sensing studies (Fu et al., 2014; Barichivich et al., 2013), as well as the satellite-based phenology depiction presented in Chapter 3 outline a significant advancement of spring phenology. The study presented in Chapter 2 depicts future spring phenology with a modelling approach and analyses species specific differences in temperature, photo-period and chilling requirements. Subsequent results proof Hypothesis 1, *"BB advancement in a warming climate will be limited by species dependent chilling and photo-period requirements"*. The implications of such a species-specific reaction on climate change are numerous, including impacts on species composition, the competitive balance between species and altered carbon exchange of ecosystems. Thus, phenological changes actually may have impacts on future climatic conditions.

Studies analysing large scale phenological shifts rely on extensive phenological datasets. Remote sensing, especially satellite-based vegetation monitoring, enables the depiction of phenology even in regions with sparse availability or absence of human observers. One of the main challenges in satellite phenology depiction is the huge number of available methods extracting phenological metrics (White et al., 2009; Fu et al., 2014). The study presented in Chapter 3 evaluates the performance of different methods, implemented and freely available in the R-package 'phenex', and outlines a disparate applicability for different compositions of land-use classes within pixels, with subsequently varying trend estimates. Consequently, the results proof Hypothesis 2, *"Methods extracting spring phenology from satellite data vary in performance depending on small scale landscape composition and thus estimate varying trends"*. This implicates a benefit of satellite sensors with high spatial resolution, thus a reduced amount of land-use classes integrated within a pixel, and the necessity of satellite product validation sites situated on different land-use classes.

Satellite product validation aims at improving the understanding of satellite-based products by providing insights into vegetational and non-vegetational factors influencing the measured signal (Soudani et al., 2008; Eklundh et al., 2011). Consequently, optical sensors at ground-level are crucial in order to bridge the gap between vegetation processes and the remotely sensed signal (Fu et al., 2014; Gamon, 2015). The study presented in Chapter 4 deals with the comparison of ground- and satellite-level NDVI products. Within the study, different factors influencing the depiction of phenology products are detected, including spatial and temporal mismatches, cloud and snow cover, illumination conditions and quality of satellite data correction algorithms (Lange et al., 2017). Further, using new satellite generation data with high spatial resolution, here from Sentinel-2, provides a benefit for phenology product depiction compared to satellite sensors with coarse spatial resolution (Munyati and Mboweni, 2013; Lange et al., 2017). Nevertheless, even with increased spatial and temporal resolution, non-vegetational effects may influence the satellite-based product (Lange et al., 2017). Consequently, the results proof Hypothesis 3, *"New satellite generations, despite high temporal and spatial resolution, benefit from near-surface sensor-derived data by the provision of outlier detection and information on environmental conditions"*.

Operating and maintaining a satellite product validation site requires high efforts. Unified measurement protocols are needed to ensure data quality and inter-comparability between different sites (Balzarolo et al., 2011; Pacheco-Labrador and Martín, 2015; Porcar-Castell et al., 2015). One of the main challenges in the context of data quality is accurate calibration of optical sensor systems, including temporal calibration stability. Sensors in long-

term continuous outdoor set-ups face accelerated degradation and subsequently calibration shifts (Porcar-Castell et al., 2015). The study presented in Chapter 5 compares a typically used campaign-based calibration with a continuous calibration utilizing redundant measurements. Results show a clear benefit of the continuous calibration approach, since the campaign-based calibration procedure can be not enough to either detect certain problems or continuously correct the measurements if necessary. Consequently, Hypothesis 4, "*Continuous calibration of unattended multi- and hyperspectral measurements is needed instead of campaign-based calibration procedures in order to acquire reliable measurements*", is proofed. Additionally, the high efforts required to calibrate unattended measurements systems can be reduced by the use of the presented approach. Thus, it is highly beneficial for the operation of a satellite product validation site.

6.6 Conclusions

The presented thesis deals with the main challenges faced in modern phenology research. Forecasting phenology, thus phenological modelling, is a timely challenge given the necessity to predict the impact of global warming on wild-growing species and agricultural crops. The modelling approach depicts an advancement of bud burst in the next centuries as long as species specific chilling and photo-period requirements are fulfilled. Assessing the present state of vegetation, thus phenological monitoring, is essential to update and validate model results. Different methods extracting spring phenology from satellite data are tested and show a varying performance depending on small scale landscape composition, with subsequently varying trend estimates of spring phenology. Further, an improved comprehension of the relationships between plant phenology and remotely sensed products is crucial to interpret satellite products. The presented satellite product validation study highlights the benefits of a validation site even when using modern satellite generation data with high temporal and spatial resolution and improved quality. Near-surface sensor-derived data provide insights into vegetational and non-vegetational factors influencing the signal measured by satellite-borne sensors. Therefore, accurate calibration of optical sensors installed at validation sites is crucial. Continuous calibration of unattended multi- and hyperspectral measurements is needed instead of campaign-based calibration procedures in order to acquire reliable measurements.

Bibliography

- Anselin L (1995) Local Indicators of Spatial Association - Lisa. *Geographical Analysis* 27(2):93–115, DOI 10.1111/j.1538-4632.1995.tb00338.x
- Ardia D, Boudt K, Carl P, Mullen KM, Peterson BG (2011) Differential evolution with DEoptim: An application to non-convex portfolio optimization. *The R Journal* 3(1):27–34
- Arnold CY (1959) The determination and significance of the base temperature in a linear heat unit system. *Proceedings of the American Society for Horticultural Science* 74:430–445
- Badeck FW, Bondeau A, Böttcher K, Doktor D, Lucht W, Schaber J, Sitch S (2004) Responses of spring phenology to climate change. *New Phytologist* 162(2):295–309, DOI 10.1111/j.1469-8137.2004.01059.x
- Baldocchi DD (2003) Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology* 9(4):479–492, DOI 10.1046/j.1365-2486.2003.00629.x
- Balzarolo M, Anderson K, Nichol C, Rossini M, Vescovo L, Arriga N, Wohlfahrt G, Calvet JC, Carrara A, Cerasoli S, Cogliati S, Daumard F, Eklundh L, Elbers JA, Evrendilek F, Handcock RN, Kaduk J, Klumpp K, Longdoz B, Matteucci G, Meroni M, Montagnani L, Ourcival JM, Sánchez-Cañete EP, Pontailier JY, Juszczak R, Scholes B, Martín MP (2011) Ground-Based Optical Measurements at European Flux Sites: A Review of Methods, Instruments and Current Controversies. *Sensors* 11(8):7954–7981, DOI 10.3390/s110807954
- Balzarolo M, Vescovo L, Hammerle A, Gianelle D, Papale D, Tomelleri E, Wohlfahrt G (2015) The relationship between ecosystem-scale hyperspectral reflectance and CO₂ exchange in European mountain grasslands. *Biogeoscience* 12:3089–3108, DOI 10.5194/bg-12-3089-2015

- Balzarolo M, Vicca S, Nguy-Robertson AL, Bonal D, Elbers JA, Fu YH, Grünwald T, Horemans JA, Papale D, Peñuelas J, Suyker A, Veroustraete F (2016) Matching the phenology of Net Ecosystem Exchange and vegetation indices estimated with MODIS and FLUXNET in-situ observations. *Remote Sensing of Environment* 174:290–300, DOI 10.1016/j.rse.2015.12.017
- Barichivich J, Briffa KR, Myneni RB, Osborn TJ, Melvin TM, Ciais P, Piao S, Tucker C (2013) Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO₂ at high northern latitudes from 1950 to 2011. *Global Change Biology* 19(10):3167–3183, DOI 10.1111/gcb.12283
- del Barrio G, Puigdefabregas J, Sanjuan ME, Stellmes M, Ruiz A (2010) Assessment and monitoring of land condition in the Iberian Peninsula, 1989-2000. *Remote Sensing of Environment* 114(8):1817–1832, DOI 10.1016/j.rse.2010.03.009
- Basler D (2016) Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central europe. *Agricultural and Forest Meteorology* 217(1):10–21, DOI 10.1016/j.agrformet.2015.11.007
- Basler D, Körner C (2012) Photoperiod sensitivity of bud burst in 14 temperate forest tree species. *Agricultural and Forest Meteorology* 165(1):73–81, DOI 10.1016/j.agrformet.2012.06.001
- Basler D, Körner C (2014) Photoperiod and temperature responses of bud swelling and bud burst in four temperate forest tree species. *Tree Physiology* 34(4):377–388, DOI 10.1093/treephys/tpu021
- de Beurs KM, Henebry GM (2004) Trend Analysis of the Pathfinder AVHRR Land (PAL) NDVI Data for the Deserts of Central Asia. *IEEE Geoscience and Remote Sensing Letters* 1(4):282–286, DOI {10.1109/LGRS.2004.834805}
- de Beurs KM, Henebry GM (2013) *Phenology: An Integrative Environmental Science. Part VI: Applications of Phenology*, chap Vegetation Phenology in Global Change Studies, pp 483–502. In: Schwartz (2013d), DOI 10.1007/978-94-007-6925-0_26
- Bluth GJS, Doiron SD, Schnetzler CC, Krueger AJ, Walter LS (1992) Global tracking of the SO₂ clouds from the June, 1991 Mount Pinatubo eruptions. *Geophysical Research Letters* 19(2):151–154, DOI {10.1029/91GL02792}

- Bogena H, Zacharias S, Papen H, Kaiser K (2011) Web page hosted at Forschungszentrum Jülich, Jülich, Germany, available online: http://teodoor.icg.kfa-juelich.de/overview-en?set_language=en (accessed 21 March 2017)
- Bonhomme R (2000) Bases and limits to using 'degree.day' units. *European Journal of Agronomy* 13(1):1–10, DOI 10.1016/S1161-0301(00)00058-7
- Botta A, Viovy N, Ciais P, Friedlingstein P, Monfray P (2000) A global prognostic scheme of leaf onset using satellite data. *Global Change Biology* 6(7):709–725, DOI 10.1046/j.1365-2486.2000.00362.x
- Bundesamt für Kartographie und Geodäsie (2012) CORINE Land Cover 10 ha. GeoBasis-DE. Geodaten der deutschen Landesvermessung. Bundesamt für Kartographie und Geodäsie, available online: <https://www.geodatenzentrum.de> (accessed on 06 Dec 2017)
- Bundesamt für Kartographie und Geodäsie (2015) Digitales Basis-Landschaftsmodell (AAA-Modellierung). GeoBasis-DE. Geodaten der deutschen Landesvermessung. Bundesamt für Kartographie und Geodäsie
- Bundesamt für Kartographie und Geodäsie (2017) Verwaltungsgebiete 1:250 000 - Stand 01.01.2017. GeoBasis-DE. Geodaten der deutschen Landesvermessung. Bundesamt für Kartographie und Geodäsie, available online: <https://www.geodatenzentrum.de> (accessed on 06 Dec 2017)
- Caffarra A, Donnelly A, Chuine I (2011) Modelling the timing of *Betula pubescens* budburst. II. Integrating complex effects of photoperiod into process-based models. *Climate Research* 46(2):159–170, DOI 10.3354/cr00983
- Chen J, Jonsson P, Tamura M, Gu ZH, Matsushita B, Eklundh L (2004) A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment* 91(3-4):332–344, DOI 10.1016/j.rse.2004.03.014
- Chen X, Wang D, Chen J, Wang C, Shen M (2018) The mixed pixel effect in land surface phenology: A simulation study. *Remote Sensing of Environment* 211:338–344, DOI 10.1016/j.rse.2018.04.030

- Chmielewski FM (2013) *Phenology: An Integrative Environmental Science. Part VI: Applications of Phenology*, chap Phenology in Agriculture and Horticulture, pp 539–561. In: Schwartz (2013d), DOI 10.1007/978-94-007-6925-0_29
- Chmielewski FM, Rötzer T (2001) Response of tree phenology to climate change across Europe. *Agricultural and Forest Meteorology* 108(2):101–112, DOI 10.1016/S0168-1923(01)00233-7
- Chmielewski FM, Heider S, Moryson S, Bruns E (2013) *Phenology: An Integrative Environmental Science. Part I: Phenological Data, Networks, and Research*, chap International Phenological Observation Networks: Concept of IPG and GPM, pp 127–153. In: Schwartz (2013a), DOI 10.1007/978-94-007-6925-0_8
- Chuine I (2000) A unified model for budburst of trees. *Journal of Theoretical Biology* 207(3):337–347, DOI 10.1006/jtbi.2000.2178
- Chuine I, Morin X, Bugmann H (2010) Warming, photoperiods, and tree phenology. *Science* 329(5989):277–278, DOI 10.1126/science.329.5989.277-e
- Chuine I, García de Cortázar-Atauri I, Kramer K, Hänninen H (2013) *Phenology: An Integrative Environmental Science. Part III: Phenological Models and Techniques*, chap Plant Development Models, pp 275–293. In: Schwartz (2013b), DOI 10.1007/978-94-007-6925-0_15
- Cogliati S, Rossini M, Julitta T, Meroni M, Schickling A, Burkart A, Pinto F, Rascher U, Colombo R (2015) Continuous and long-term measurements of reflectance and sun-induced chlorophyll fluorescence by using novel automated field spectroscopy systems. *Remote Sensing of Environment* 164:270–281, DOI 10.1016/j.rse.2015.03.027
- Cong N, Piao S, Chen A, Wang X, Lin X, Chen S, Han S, Zhou G, Zhang X (2012) Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agricultural and Forest Meteorology* 165(Supplement C):104–113, DOI 10.1016/j.agrformet.2012.06.009
- Cong N, Wang T, Nan H, Ma Y, Wang X, Myneni RB, Piao S (2013) Changes in satellite-derived spring vegetation green-up date and its linkage to climate in China from 1982 to 2010: a multimethod analysis. *Global Change Biology* 19(3):881–891, DOI {10.1111/gcb.12077}

- Copernicus Sentinel data (2017) *Scientific Data Hub*. Available online: <http://scihub.copernicus.eu> (accessed 21 December 2016)
- Cracknell AP (1997) *The Advanced Very High Resolution Radiometer*. Taylor & Francis, London, Bristol (PA), ISBN:9780748402090
- Damm A, Guanter L, Verhoef W, Schläpfer D, Garbari S, Schaepman ME (2015) Impact of varying irradiance on vegetation indices and chlorophyll fluorescence derived from spectroscopy data. *Remote Sensing of Environment* 156:202–215, DOI 10.1016/j.rse.2014.09.031
- Daumard F, Champagne S, Fournier A, Goulas Y, Ounis A, Hanocq JF, Moya I (2010) A field platform for continuous measurement of canopy fluorescence. *IEEE Transactions on Geoscience and Remote Sensing* 48:3358–3368, DOI 10.1109/TGRS.2010.2046420
- Dechant B, Lange M, Cuntz M, Doktor D (2017) Continuous and long-term monitoring of the radiometric calibration of unattended multi- and hyperspectral measurements of vegetation canopy optical properties., manuscript in preparation
- Dechant B, Lange M, Cuntz M, Doktor D (2018) Extremely high correlation of solar irradiance in the visible and near-infrared: implications for multispectral vegetation reflectance monitoring., manuscript in preparation
- Delbart N, Kergoat L, Le Toan T, Lhermitte J, Picard G (2005) Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment* 97(1):26–38, DOI 10.1016/j.rse.2005.03.011
- Delbart N, Le Toan T, Kergoat L, Fedotova V (2006) Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982-2004). *Remote Sensing of Environment* 101(1):52–62, DOI 10.1016/j.rse.2005.11.012
- Dierenbach J, Badeck FW, Schaber J (2013) The plant phenological online database (PPODB): an online database for long-term phenological data. *International Journal of Biometeorology* 57(5):1–8, DOI 10.1007/s00484-013-0650-2
- DLR (2018) TIMELINE - TIME Series Processing of Medium Resolution Earth Observation Data assessing Long -Term Dynamics In our Natural Environment.

- Available online: https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9036/15754_read-38904/ (accessed on 30 Aug 2018)
- Doktor D (2008) Using satellite imagery and ground observations to quantify the effect of intra-annually changing temperature patterns on spring time phenology. PhD thesis, University of London, London, UK
- Doktor D, Lange M (2017) Disparate applicability and broad spatio-temporal satellite resolution affects extracted trends of European spring phenology for 1989-2007, manuscript under revision
- Doktor D, Bondeau A, Koslowski D, Badeck FW (2009) Influence of heterogeneous landscapes on computed green-up dates based on daily AVHRR NDVI observations. *Remote Sensing of Environment* 113(12):2618–2632, DOI 10.1016/j.rse.2009.07.020
- Doktor D, Lausch A, Sprengler D, Thurner M (2014) Extraction of Plant Physiological Status from Hyperspectral Signatures Using Machine Learning Methods. *Remote Sensing* 6(12):12,247–12,274, DOI 10.3390/rs61212247
- Donohue RJ, Roderick ML, McVicar TR, Farquhar GD (2013) Impact of CO₂ fertilization on maximum foliage cover across the globe's warm, arid environments. *Geophysical Research Letters* 40(12):3031–3035, DOI {10.1002/grl.50563}
- Drolet G, Wade T, Nichol CJ, MacLellan C, Levula J, Porcar-Castell A, Nikinmaa E, Vesala T (2014) A temperature-controlled spectrometer system for continuous and unattended measurements of canopy spectral radiance and reflectance. *International Journal of Remote Sensing* 35(5):1769–1785, DOI 10.1080/01431161.2014.882035
- Duchemin B, Goubier J, Courrier G (1999) Monitoring phenological key stages and cycle duration of temperate deciduous forest ecosystems with NOAA/AVHRR data. *Remote Sensing of Environment* 67(1):68–82, DOI 10.1016/S0034-4257(98)00067-4
- Eastman JR, Sangermano F, Machado EA, Rogan J, Anyamba A (2013) Global Trends in Seasonality of Normalized Difference Vegetation Index (NDVI), 1982-2011. *Remote Sensing* 5(10):4799–4818, DOI {10.3390/rs5104799}
- Eidenshink J (2006) A 16-year time series of 1 km AVHRR satellite data of the conterminous United States and Alaska. *Photogrammetric Engineering and Remote Sensing* 72(9):1027–1035, DOI {10.14358/PERS.72.9.1027}

- Eklundh L, Jin H, Schubert P, Guzinski R, Heliasz M (2011) An Optical Sensor Network for Vegetation Phenology Monitoring and Satellite Data Calibration. *Sensors* 11(8):7678, DOI 10.3390/s110807678
- Enke W, Kreienkamp F (2006) WETTREG A1B SCENARIO RUN, UBA PROJECT. Tech. rep., World Data Center for Climate, 10-year slices, e.g. CERA-DB "WR_A1B_2071_2080", http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=WR_A1B_2071_2080; Additionally, Scenarios A2 and B1 as well as wet and dry realisations were used.
- Estrella N, Sparks TH, Menzel A (2009) Effects of temperature, phase type and timing, location, and human density on plant phenological responses in Europe. *Climate Research* 39(3):235–248, DOI 10.3354/cr00818
- European Environment Agency (2013) CORINE Land Cover 2000 raster data. Available online: <http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-raster-4> (accessed on 17 May 2019)
- Fischer A (1994) A simple model for the temporal variations of NDVI at regional scale over agricultural countries. Validation with ground radiometric measurements. *International Journal of Remote Sensing* 15(7):1421–1446, DOI 10.1080/01431169408954175
- Fisher JJ, Mustard JF (2007) Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment* 109(3):261–273, DOI 10.1016/j.rse.2007.01.004
- Fisher JJ, Mustard JF, Vadeboncoeur MA (2006) Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment* 100(2):265–279, DOI 10.1016/j.rse.2005.10.022
- Fitter AH, Fitter RSR (2002) Rapid changes in flowering time in British plants. *Science* 296(5573):1689–1691, DOI 10.1126/science.1071617
- Fletcher K (2012) SENTINEL-2: ESA's Optical High-Resolution Mission for GMES Operational Services (ESA SP-1322/2). Tech. rep., European Space Agency, available online: http://esamultimedia.esa.int/multimedia/publications/SP-1322_2/ (accessed on 13 Oct 2017)

- FLUXNET (2017) Web Page. Lawrence Berkeley National Laboratory, CA, USA, available online: <http://fluxnet.fluxdata.org/> (accessed on 10 Nov 2017)
- Friedrich K, Koslowsky D (2009) *Inter-comparison of MEDOKADS and NOAA/NASA pathfinder AVHRR land NDVI time series*, Taylor & Francis, London, pp 103–116
- Fu Y, Campioli M, Van Oijen M, Deckmyn G, Janssens I (2012) Bayesian comparison of six different temperature-based budburst models for four temperate tree species. *Ecological Modelling* 230(1):92–100, DOI 10.1016/j.ecolmodel.2012.01.010
- Fu Y, Piao S, Op de Beeck M, Cong N, Zhao H, Zhang Y, Menzel A, Janssens I (2014) Recent spring phenology shifts in western Central Europe based on multiscale observations. *Global Ecology and Biogeography* 23(11):1255–1263, DOI 10.1111/geb.12210
- Fu Y, Piao S, Zhao H, Jeong SJ, Wang X, Vitasse Y, Ciais P, Janssens I (2014) Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern middle and high latitudes. *Global Change Biology* 20(12):3743–3755, DOI 10.1111/gcb.12610
- Fu Y, Piao S, Vitasse Y, Zhao H, De Boeck H, Liu Q, Yang H, Weber U, Hänninen H, Janssens I (2015a) Increased heat requirement for leaf flushing in temperate woody species over 1980–2012: effects of chilling, precipitation and insolation. *Global Change Biology* 21(7):2687–2697, DOI 10.1111/gcb.12863
- Fu Y, Zhao H, Piao S, Peaucelle M, Peng S, Zhou G, Ciais P, Huang M, Menzel A, Penuelas J, Song Y, Vitasse Y, Zeng Z, Janssens I (2015b) Declining global warming effects on the phenology of spring leaf unfolding. *Nature* 526(1):104–107, DOI 10.1038/nature15402
- Füssel HM, Jol A, Hildén M, Christiansen T, Kurnik B, Hemming D, Hartley A, Lowe J, Meiner A, Kristensen P, Vanneuville W, Camia A, San-Miguel-Ayanz J, Bastrup-Birk A, Marx A, Jones A, Cherlet M, Louwagie G, Jarosinska D, Watkiss P (2012) *Climate change, impacts and vulnerability in Europe 2012*. European Environment Agency, ISBN: 978-92-9213-346-7
- Gamon JA (2015) Reviews and Syntheses: optical sampling of the flux tower footprint. *Biogeosciences* 12(14):4509–4523, DOI 10.5194/bg-12-4509-2015

- Gamon JA, Cheng Y, Claudio H, Mackinney L, Sims D (2006a) A mobile tram system for systematic sampling of ecosystem optical properties. *Remote Sensing of Environment* 103(3):246–254, DOI 10.1016/j.rse.2006.04.006
- Gamon JA, Rahman AF, Dungan JL, Schildhauer M, Huemmrich KF (2006b) Spectral Network (SpecNet) - What is it and why do we need it? *Remote Sensing of Environment* 103(3):227–235, DOI 10.1016/j.rse.2006.04.003
- Gamon JA, Coburn C, Flanagan LB, Huemmrich KF, Kiddle C, Sanchez-Azofeifa GA, Thayer DR, Vescovo L, Gianelle D, Sims DA, Rahman AF, Pastorello GZ (2010) SpecNet revisited: bridging flux and remote sensing communities. *Canadian Journal of Remote Sensing* 36(sup2):376–390, DOI 10.5589/m10-067
- Garonna I, De Jong R, De Wit AJW, Mùcher CA, Schmid B, Schaepman ME (2014) Strong contribution of autumn phenology to changes in satellite-derived growing season length estimates across Europe (1982-2011). *Global Change Biology* 20(11):3457–3470, DOI 10.1111/gcb.12625
- Geng L, Ma M, Wang H (2016) An Effective Compound Algorithm for Reconstructing MODIS NDVI Time Series Data and Its Validation Based on Ground Measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9(8):3588–3597, DOI 10.1109/JSTARS.2015.2495112
- Gitelson AA, Gritz Y, Merzlyak MN (2003) Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology* 160(3):271–282, DOI 10.1078/0176-1617-00887
- Gonsamo A, Chen JM, Price DT, Kurz WA, Wu C (2012) Land surface phenology from optical satellite measurement and CO2 eddy covariance technique. *Journal of Geophysical Research - Biogeosciences* 117, DOI {10.1029/2012JG002070}
- Guanter L, Kaufmann H, Segl K, Foerster S, Rogass C, Chabrillat S, Kuester T, Hollstein A, Rossner G, Chlebek C, Straif C, Fischer S, Schrader S, Storch T, Heiden U, Mueller A, Bachmann M, Mühle H, Müller R, Habermeyer M, Ohndorf A, Hill J, Buddenbaum H, Hostert P, van der Linden S, Leitão P, Rabe A, Doerffer R, Krasemann H, Xi H, Mauser W, Hank T, Locherer M, Rast M, Staenz K, Sang B (2015) The EnMAP Spaceborne

Bibliography

- Imaging Spectroscopy Mission for Earth Observation. *Remote Sensing* 7(7):8830–8857, DOI 10.3390/rs70708830
- Hamunyela E, Verbesselt J, Roerink G, Herold M (2013) Trends in Spring Phenology of Western European Deciduous Forests. *Remote Sensing* 5(12):6159–6179, DOI {10.3390/rs5126159}
- Hänninen H (1990) Modelling bud dormancy release in trees from cool and temperate regions. *Acta Forestalia Fennica* 213:1–47, DOI 10.14214/aff.7660
- Hänninen H (1991) Does climatic warming increase the risk of frost damage in northern trees. *Plant Cell and Environment* 14(5):449–454, DOI 10.1111/j.1365-3040.1991.tb01514.x
- Hänninen H, Kramer K (2007) A framework for modelling the annual cycle of trees in boreal and temperate regions. *Silva Fennica* 41(1):167–205, DOI 10.14214/sf.313
- Hänninen H, Tanino K (2011) Tree seasonality in a warming climate. *Trends in Plant Science* 16(8):412–416, DOI 10.1016/j.tplants.2011.05.001
- Hansen J, Ruedy R, Glascoe J, Sato M (1999) GISS analysis of surface temperature change. *Journal of Geophysical Research: Atmospheres* 104(D24):30,997–31,022, DOI 10.1029/1999JD900835
- Hansen J, Sato M, Ruedy R, Lo K, Lea DW, Medina-Elizade M (2006) Global temperature change. *Proceedings of the National Academy of Sciences* 103(39):14,288–14,293, DOI 10.1073/pnas.0606291103
- Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. *Reviews of Geophysics* 48(4), DOI 10.1029/2010RG000345
- Heide OM (1993) Daylength and thermal time responses of budburst during dormancy release in some northern deciduous trees. *Physiologia Plantarum* 88(4):531–540, DOI 10.1111/j.1399-3054.1993.tb01368.x
- Helman D (2018) Land surface phenology: What do we really 'see' from space? *Science of The Total Environment* 618:665–673, DOI 10.1016/j.scitotenv.2017.07.237

- Henebry GM, de Beurs KM (2013) *Phenology: An Integrative Environmental Science. Part IV: Sensor-Derived Phenology*, chap Remote Sensing of Land Surface Phenology: A Prospectus, pp 385–411. In: Schwartz (2013c), DOI 10.1007/978-94-007-6925-0_21
- Hermance JF, Jacob RW, Bradley BA, Mustard JF (2007) Extracting phenological signals from multiyear AVHRR NDVI time series: Framework for applying high-order annual splines with roughness damping. *IEEE Transactions on Geoscience and Remote Sensing* 45(10):3264–3276, DOI 10.1109/TGRS.2007.903044
- Heumann BW, Seaquist JW, Eklundh L, Jönsson P (2007) AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982-2005. *Remote Sensing of Environment* 108:385–392, DOI 10.1016/j.rse.2006.11.025
- Hilker T, Coops NC, Nesic Z, Wulder MA, Black AT (2007) Instrumentation and approach for unattended year round tower based measurements of spectral reflectance. *Computers and Electronics in Agriculture* 56(1):72–84, DOI 10.1016/j.compag.2007.01.003
- Hilker T, Coops NC, Hall FG, Black AT, Chen B, Krishnan P, Wulder MA, Sellers PJ, Middleton EM, Huemmrich KF (2008a) A modeling approach for upscaling gross ecosystem production to the landscape scale using remote sensing data. *Journal of Geophysical Research: Biogeosciences* 113(G3), DOI 10.1029/2007JG000666
- Hilker T, Coops NC, Hall FG, Black AT, Wulder MA, Nesic Z, Krishnan P (2008b) Separating physiologically and directionally induced changes in PRI using BRDF models. *Remote Sensing of Environment* 112(6):2777–2788, DOI 10.1016/j.rse.2008.01.011
- Hilker T, Hall FG, Coops NC, Lyapustin A, Wang Y, Nesic Z, Grant N, Black AT, Wulder MA, Kljun N (2010) Remote sensing of photosynthetic light-use efficiency across two forested biomes: Spatial scaling. *Remote Sensing of Environment* 114(12):2863–2874, DOI 10.1016/j.rse.2010.07.004
- Hilker T, Hall FG, Coops NC, Collatz JG, Black AT, Tucker CJ, Sellers PJ, Grant N (2013) Remote sensing of transpiration and heat fluxes using multi-angle observations. *Remote Sensing of Environment* 137(1):31–42, DOI 10.1016/j.rse.2013.05.023
- Hilker T, Hall FG, Coops NC, Black AT, Jassal R, Mathys A, Grant N (2014) Potentials and limitations for estimating daytime ecosystem respiration by combining tower-based re-

- mote sensing and carbon flux measurements. *Remote Sensing of Environment* 150(1):44–52, DOI 10.1016/j.rse.2014.04.018
- Hill J, Stellmes M, Udelhoven T, Röder A, Sommer S (2008) Mediterranean desertification and land degradation: Mapping related land use change syndromes based on satellite observations. *Global and Planetary Change* 64(3-4):146–157, DOI 10.1016/j.gloplacha.2008.10.005
- Hmimina G, Dufrêne E, Pontailier JY, Delpierre N, Aubinet M, Caquet B, de Grandcourt A, Burban B, Flechard C, Granier A, Gross P, Heinesch B, Longdoz B, Moureaux C, Ourcival JM, Rambal S, Saint André L, Soudani K (2013) Evaluation of the potential of modis satellite data to predict vegetation phenology in different biomes: An investigation using ground-based ndvi measurements. *Remote Sensing of Environment* 132:145–158, DOI <https://doi.org/10.1016/j.rse.2013.01.010>
- Holben BN (1986) Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing* 7(11):1417–1434, DOI 10.1080/01431168608948945
- Huber S, Tagesson T, Fensholt R (2014) An automated field spectrometer system for studying VIS, NIR and SWIR anisotropy for semi-arid savanna. *Remote Sensing of Environment* 152(1):547–556, DOI 10.1016/j.rse.2014.06.007
- Huemmrich K, Black T, Jarvis P, McCaughey J, Hall F (1999) High temporal resolution NDVI phenology from micrometeorological radiation sensors. *Journal of Geophysical Research - Atmospheres* 104(D22):27,935–27,944, DOI { 10.1029/1999JD900164 }
- ICOS ERIC (2017) Web Page. ICOS ERIC Head Office, Helsinki, Finland, available online: <http://www.icos-ri.eu/> (accessed on 19 May 2017)
- Intergovernmental Panel on Climate Change (2007) *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]
- Intergovernmental Panel on Climate Change (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects*. Cambridge University Press,

- Cambridge, United Kingdom and New York, NY, USA, Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]
- Jacob D, Mahrenholz P (2006) REMO A1B SCENARIO RUN, UBA PROJECT, DATASTREAM 3. Tech. rep., World Data Center for Climate, CERA-DB "REMO_UBA_A1B_D3", http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=REMO_UBA_A1B_D3; Additionally, Scenarios A2 and B1 were used.
- Jacob D, Elizalde A, Kotova L, Pfeifer S, Moseley C, Kumar P, Rechid D, Teichmann C (2006) Regional climate modelling. <http://www.remo-rcm.de>, (accessed on 01 August 2013)
- Jacquemoud S (1993) Inversion of the PROSPECT + SAIL canopy reflectance model from AVIRIS equivalent spectra: Theoretical study. *Remote Sensing of Environment* 44:281–292, DOI 10.1016/0034-4257(93)90022-P
- Jacquemoud S, Verhoef W, Baret F, Bacour C, Zarco-Tejada PJ, Asner GP, François C, Ustin SL (2009) PROSPECT plus SAIL models: A review of use for vegetation characterization. *Remote Sensing Of Environment* 113:S56–S66, DOI 10.1016/j.rse.2008.01.026
- Jeong SJ, Ho CH, Gim HJ, Brown ME (2011) Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. *Global Change Biology* 17(7):2385–2399, DOI 10.1111/j.1365-2486.2011.02397.x
- Jiang Z, Huete AR, Chen J, Chen Y, Li J, Yan G, Zhang X (2006) Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment* 101(3):366–378, DOI 10.1016/j.rse.2006.01.003
- Jin H, Eklundh L (2014) A physically based vegetation index for improved monitoring of plant phenology. *Remote Sensing of Environment* 152:512–525, DOI 10.1016/j.rse.2014.07.010
- Jin H, Eklundh L (2015) In situ calibration of light sensors for long-term monitoring of vegetation. *IEEE Transactions on Geoscience and Remote Sensing* 53(6):3405–3416, DOI 10.1109/TGRS.2014.2375381

Bibliography

- Johnson I, Thornley JHM (1985) Temperature dependence of plant and crop processes. *Annals of Botany* 55(1):1–24, DOI 10.1093/oxfordjournals.aob.a086868
- Jönsson AM, Eklundh L, Hellström M, Barring L, Jonsson P (2010) Annual changes in MODIS vegetation indices of swedish coniferous forests in relation to snow dynamics and tree phenology. *Remote Sensing of Environment* 114(11):2719–2730, DOI 10.1016/j.rse.2010.06.005
- Julien Y, Sobrino JA (2009) Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing* 30(13):3495–3513, DOI {10.1080/01431160802562255}
- Justice C, Belward A, Morisette J, Lewis P, Privette J, Baret F (2000) Developments in the 'validation' of satellite sensor products for the study of the land surface. *International Journal of Remote Sensing* 21(17):3383–3390, DOI 10.1080/014311600750020000
- Karlsen SR, Tolvanen A, Kubin E, Poikolainen J, Hännikä KA, Johansen B, Danks FS, Aspholm P, Wielgolaski FE, Makarova O (2008) MODIS-NDVI-based mapping of the length of the growing season in northern Fennoscandia. *International Journal of Applied Earth Observation and Geoinformation* 10(3):253–266, DOI 10.1016/j.jag.2007.10.005
- Keeling CD, Chin JFS, Whorf TP (1996) Increased activity of northern vegetation inferred from atmospheric CO₂ measurements. *Nature* 382(6587):146–149, DOI 10.1038/382146a0
- Keenan T (2015) Phenology: Spring greening in a warming world. *Nature* 526(7571):48–49, DOI 10.1038/nature15633
- Kidwell KB (1995) NOAA Polar Orbiter Data (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA-11, NOAA-12, and NOAA-14) Users Guide. Tech. rep., U.S. Department of Commerce; National Oceanic and Atmospheric Administration; National Environmental Satellite, Data and Information Service; National Climatic Data Center; Satellite Data Service Division, Washington, DC, USA
- Köppen W (1918) Klassifikation der Klimate nach Temperatur, Niederschlag und Jahresablauf (Classification of climates according to temperature, precipitation and seasonal cycle). *Petermanns Geogr Mitt* 64(1):193–203, 243–248

- Koslowsky D, Billing H, Eckardt M (2001) *Sensor degradation and intercalibration of the short wave channels of the AVHRR-NOAA 11/14/16 satellites*. In: Proceedings of the 2001 EUMETSAT Meteorological Satellite Data Users' Conference, pp 107–113
- Kramer K (1994) Selecting a model to predict the onset of growth of *fagus sylvatica*. *Journal of Applied Ecology* 31:172–181, DOI 10.2307/2404609
- Kramer K (1995) Modelling comparison to evaluate the importance of phenology for the effects of climate change on growth of temperate-zone deciduous trees. *Climate Research* 5:119–130, DOI 10.3354/cr005119
- Kramer K, Friend A, Leinonen I (1996) Modelling comparison to evaluate the importance of phenology and spring frost damage for the effects of climate change on growth of mixed temperate-zone deciduous forests. *Climate Research* 7:31–41, DOI 10.3354/cr007031
- Kunstmann H, Schneider K, Forkel R, Knoche R (2004) Impact analysis of climate change for an Alpine catchment using high resolution dynamic downscaling of ECHAM4 time slices. *Hydrology and Earth System Sciences* 8(6):1030 – 1044, DOI 10.5194/hess-8-1031-2004
- Landsberg JJ (1974) Apple fruit bud development and growth; analysis and an empirical model. *Annals of Botany* 38(158):1013–1023
- Lange M (2013) *phenmod: Auxiliary functions for phenological data processing, modelling and result handling*. R package version 1.2-2, <http://CRAN.R-project.org/package=phenmod> (accessed on 20 August 2013)
- Lange M, Doktor D (2017) *phenex: Auxiliary Functions for Phenological Data Analysis*. R package version 1.4-5, <https://CRAN.R-project.org/package=phenex> (accessed on 29 May 2017)
- Lange M, Schaber J, Marx A, Jäckel G, Badeck FW, Seppelt R, Doktor D (2016) Simulation of forest tree species' bud burst dates for different climate scenarios: chilling requirements and photo-period may limit bud burst advancement. *International Journal of Biometeorology* 60(11):1711–1726, DOI 10.1007/s00484-016-1161-8
- Lange M, Dechant B, Rebmann C, Vohland M, Cuntz M, Doktor D (2017) Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors. *Sensors* 17(8), DOI 10.3390/s17081855

Bibliography

- Laube J, Sparks T, Estrella N, Höfler J, Ankerst D, Menzel A (2014) Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology* 20(1):170–182, DOI 10.1111/gcb.12360
- Lauscher F (1978) Neue Analysen ältester und neuerer phänologischer Reihen. *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B* 26(4):373–385, DOI 10.1007/BF02243239
- Leuning R, Hughes D, Daniel P, Coops N, Newnham G (2006) A multi-angle spectrometer for automatic measurement of plant canopy reflectance spectra. *Remote Sensing of Environment* 103(1):236–245, DOI 10.1016/j.rse.2005.06.016
- Liang LA, Schwartz MD, Fei SL (2011) Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sensing of Environment* 115(1):143–157, DOI 10.1016/j.rse.2010.08.013
- Linkosalo T, Carter TR, Hakkinen R, Hari P (2000) Predicting spring phenology and frost damage risk of *Betula* spp. under climatic warming: a comparison of two models. *Tree Physiology* 20(17):1175–1182, DOI 10.1093/treephys/20.17.1175
- Linkosalo T, Hakkinen R, Hanninen H (2006) Models of the spring phenology of boreal and temperate trees: is there something missing? *Tree Physiology* 26(9):1165–1172, DOI 10.1093/treephys/26.9.1165
- Lobell DB, Asner GP (2004) Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of Environment* 93(3):412–422, DOI 10.1016/j.rse.2004.08.002
- LTER (2018) Web Page. LTER Network Communications Office, NCEAS, UCSB, 735 State Street, Suite 300, Santa Barbara, CA 93101, Available online: <https://lternet.edu/> (accessed on 22 Feb 2018)
- Lüdeke MKB, Ramage PH, Kohlmaier GH (1996) The use of satellite NDVI data for the validation of global vegetation phenology models: application to the Frankfurt Biosphere Model. *Ecological Modelling* 91(1):255 – 270, DOI 10.1016/0304-3800(95)00192-1
- MacArthur A, Robinson I, Rossini M, Davis N, MacDonald K (2014) A dual-field-of-view spectrometer system for reflectance and fluorescence measurements (Piccolo Doppio) and

- correction of etaloning. Proceedings of the Fifth International Workshop on Remote Sensing of Vegetation Fluorescence, 22-24 April, European Space Agency, Paris, United Kingdom
- Maisongrande P, Duchemin B, Dedieu G (2004) VEGETATION/SPOT: an operational mission for the Earth monitoring; presentation of new standard products. *International Journal of Remote Sensing* 25(1):9–14, DOI 10.1080/0143116031000115265
- Menzel A (2005) A 500 year pheno-climatological view on the 2003 heatwave in Europe assessed by grape harvest dates. *Meteorologische Zeitschrift* 14(1):75–77, DOI 10.1127/0941-2948/2005/0014-0075
- Menzel A (2013a) *Phenology: An Integrative Environmental Science. Part I: Phenological Data, Networks, and Research*, chap Europe, pp 53–65. In: Schwartz (2013a), DOI 10.1007/978-94-007-6925-0_4
- Menzel A (2013b) *Phenology: An Integrative Environmental Science. Part III: Phenological Models and Techniques*, chap Plant Phenological "Fingerprints", pp 335–350. In: Schwartz (2013b), DOI 10.1007/978-94-007-6925-0_18
- Menzel A, Fabian P (1999) Growing season extended in Europe. *Nature* 397(1):659, DOI 10.1038/17709
- Meroni M, Rossini M, Guanter L, Alonso L, Rascher U, Colombo R, Moreno J (2009) Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment* 113(10):2037–2051, DOI 10.1016/j.rse.2009.05.003
- Meroni M, Barducci A, Cogliati S, Castagnoli F, Rossini M, Busetto L, Migliavacca M, Cremonese E, Galvagno M, Colombo R, di Cella UM (2011) The hyperspectral irradiometer, a new instrument for long-term and unattended field spectroscopy measurements. *Review of Scientific Instruments* 82(4), DOI 10.1063/1.3574360
- Meynen E, Schmithüsen J (1962) *Handbuch der naturräumlichen Gliederung Deutschlands*. Selbstverlag der Bundesanstalt für Landeskunde, Bad Godesberg
- Migliavacca M, Sonnentag O, Keenan TF, Cescatti A, O' Keefe J, Richardson AD (2012) On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model. *Biogeosciences* 9(1):2063–2083, DOI 10.5194/bg-9-2063-2012

Bibliography

- Morin X, Lechowicz MJ, Augspurger C, O' Keefe J, Viner D, Chuine I (2009) Leaf phenology in 22 North American tree species during the 21st century. *Global Change Biology* 15(4):961–975, DOI 10.1111/j.1365-2486.2008.01735.x
- Morisette JT, Baret F, Liang S (2006) Special Issue on Global Land Product Validation. *IEEE Transactions on Geoscience and Remote Sensing* 44(7):1695–1697, DOI 10.1109/TGRS.2006.877436
- Moser L, Fonti P, Buentgen U, Esper J, Luterbacher J, Franzen J, Frank D (2010) Timing and duration of European larch growing season along altitudinal gradients in the Swiss Alps. *Tree Physiology* 30(2):225–233, DOI 10.1093/treephys/tpp108
- Mueller-Wilm U, Devignot O, Pessiot L (2016) *S2 MPC Sen2Cor Configuration and User Manual*. Available online: <http://step.esa.int/thirdparties/sen2cor/2.3.1> (accessed on 01 06 2017)
- Müller-Westermeier G (2002) Klimastatusbericht 2001: Das Klima in Deutschland. Deutscher Wetterdienst, available online: https://www.dwd.de/DE/leistungen/klimastatusbericht/publikationen/ksb2001_pdf/01_2001.html?nn=16102 (accessed on 18 Oct 2017)
- Munyati C, Mboweni G (2013) Variation in NDVI values with change in spatial resolution for semi-arid savanna vegetation: a case study in northwestern South Africa. *International Journal of Remote Sensing* 34(7):2253–2267, DOI 10.1080/01431161.2012.743692
- Murray MB, Cannell MGR, Smith RI (1989) Date of budburst of fifteen tree species in Britain following climate warming. *The Journal of Applied Ecology* 26:693, DOI 10.2307/2404093
- Myking T, Heide OM (1995) Dormancy release and chilling requirement of buds of latitudinal ecotypes of *Betula pendula* and *B. pubescens*. *Tree Physiology* 15(11):697–704, DOI 10.1093/treephys/15.11.697
- Myneni RB, Keeling CD, Tucker CJ, Asrar G, Nemani RR (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 386(6626):698–702, DOI 10.1038/386698a0

- Nakaji T, Ide R, Oguma H, Saigusa N, Fujinuma Y (2007) Utility of spectral vegetation index for estimation of gross co₂ flux under varied sky conditions. *Remote Sensing of Environment* 109(3):274–284, DOI 10.1016/j.rse.2007.01.006
- Nakićenović N, Alcamo J, Davis D, de Vries B, Fenhann J, Gaffin S, Gregory K, Grubler A, Jung TY, Kram T, Lebre La Rovere E, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, , Dadi Z (2000) Special report on emissions scenarios. Cambridge University Press
- NASA (2017) MODIS Web Page. National Aeronautics and Space Administration, NASA Headquarters, Washington, DC, USA; Available online: <https://modis.gsfc.nasa.gov/> (accessed on 13 Oct 2017)
- Nestola E, Calfapietra C, Emmerton CA, Wong CYS, Thayer DR, Gamon JA (2016) Monitoring Grassland Seasonal Carbon Dynamics, by Integrating MODIS NDVI, Proximal Optical Sampling, and Eddy Covariance Measurements. *Remote Sensing* 8(3):260, DOI {10.3390/rs8030260}
- Nicodemus FE, Richmond JC, Hsia JJ, Ginsberg IW, Limperis T (1977) *Geometrical Considerations and Nomenclature for Reflectance*. NBS monograph, U.S. Government Printing Office: Washington, D.C., USA, ISBN:0-86720-294-7
- NOAA (2017) NOAA satellite information system. Available online: <http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html> (accessed on 09 Oct 2017)
- Olsson C, Jönsson A (2014) Process-based models not always better than empirical models for simulating budburst of Norway spruce and birch in europe. *Global Change Biology* 20(11):3492–3507, DOI 10.1111/gcb.12593
- Pacheco-Labrador J, Martín MP (2015) Characterization of a Field Spectroradiometer for Unattended Vegetation Monitoring. Key Sensor Models and Impacts on Reflectance. *Sensors* 15(2):4154–4175, DOI 10.3390/s150204154
- Pacheco-Labrador J, Martín MP, Riaño D, Hilker T, Carrara A (2016) New approaches in multi-angular proximal sensing of vegetation: Accounting for spatial heterogeneity and diffuse radiation in directional reflectance distribution models. *Remote Sensing of Environment* 187:447–457, DOI 10.1016/j.rse.2016.10.051

Bibliography

- Park IW (2016) Timing the bloom season: a novel approach to evaluating reproductive phenology across distinct regional flora. *Landscape Ecology* 31(7):1567–1579", DOI 10.1007/s10980-016-0339-0
- Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology* 13(1):1860–1872, DOI 10.1111/j.1365-2486.2007.01404.x
- Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421(6918):37–42, DOI 10.1038/nature01286
- Paschotta R (2013) Tutorial on "Passive Fiber Optics". Tech. rep., RP Photonics, available online: https://www.rp-photonics.com/passive_fiber_optics.html, (accessed on 16 Jul 2018)
- Peñuelas J, Rutishauser T, Filella I (2009) Phenology feedbacks on climate change. *Science* 324:887–888, DOI 10.1126/science.1173004
- Peñuelas J, Sardans J, Estiarte M, Ogaya R, Carnicer J, Coll M, Barbeta A, Rivas-Ubach A, Llusia J, Garbulsky M, Filella I, Jump A (2013) Evidence of current impact of climate change on life: a walk from genes to the biosphere. *Global Change Biology* 19(8):2303–2338, DOI 10.1111/gcb.12143
- Pearson K (1920) Notes on the History of Correlation. *Biometrika* 13(1):25–45, DOI 10.1093/biomet/13.1.25
- Peel M, Finlayson B, McMahon T (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* 11(5):1633–1644, DOI 10.5194/hess-11-1633-2007
- Piao S, Cui M, Chen A, Wang X, Ciais P, Liu J, Tang Y (2011) Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau. *Agricultural and Forest Meteorology* 151(12):1599–1608, DOI 10.1016/j.agrformet.2011.06.016
- Pinzon JE, Tucker CJ (2014) A Non-Stationary 1981-2012 AVHRR NDVI3g Time Series. *Remote Sensing* 6(8):6929–6960, DOI { 10.3390/rs6086929 }

- Pope K, Dose V, Da Silva D, Brown P, Leslie C, Dejong T (2013) Detecting nonlinear response of spring phenology to climate change by Bayesian analysis. *Global Change Biology* 19(5):1518–1525, DOI 10.1111/gcb.12130
- Porcar-Castell A, Tyystjarvi E, Atherton J, van der Tol C, Flexas J, Pfundel EE, Moreno J, Frankenberg C, Berry JA (2014) Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *Journal of Experimental Botany* 65(15):4065–4095, DOI 10.1093/jxb/eru191
- Porcar-Castell A, Mac Arthur A, Rossini M, Eklundh L, Pacheco-Labrador J, Anderson K, Balzarolo M, Martín MP, Jin H, Tomelleri E, Cerasoli S, Sakowska K, Hueni A, Julitta T, Nichol CJ, Vescovo L (2015) EUROSPEC: at the interface between remote-sensing and ecosystem CO₂ flux measurements in Europe. *Biogeosciences* 12(20):6103–6124, DOI 10.5194/bg-12-6103-2015
- Press WH (1992) *Numerical recipes in C: The Art of Scientific Computing*, vol 1, 2nd edn. Cambridge University Press, Cambridge
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, available online: <http://www.R-project.org/> (accessed on 01 06 2017)
- Réaumur M (1735) *Observations du thermomètres, faites à Paris pendant l'année 1735, comparées avec celles qui ont été faites sous la ligne, à l'isle de France, à Alger et quelques unes de nos isles de l'Amérique*. Académie des Sciences de Paris, 545
- Reed BC (2006) Trend Analysis of Time-Series Phenology of North America Derived from Satellite Data. *GIScience & Remote Sensing* 43(1):24–38, DOI 10.2747/1548-1603.43.1.24
- Reed BC, Brown JF, VanderZee D, Loveland TR, Merchant JW, Ohlen DO (1994) Measuring phenological variability from satellite imagery. *Journal of Vegetation Science* 5(5):703–714, DOI 10.2307/3235884
- Repo T, Hanninen H, Kellomaki S (1996) The effects of long-term elevation of air temperature and CO₂ on the frost hardiness of Scots pine. *Plant Cell and Environment* 19(2):209–216, DOI 10.1111/j.1365-3040.1996.tb00242.x

Bibliography

- Richardson AD, Jenkins JP, Braswell BH, Hollinger DY, Ollinger SV, Smith ML (2007) Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia* 152(2):323–334, DOI 10.1007/s00442-006-0657-z
- Richardson AD, Klosterman S, Toomey M (2013) *Phenology: An Integrative Environmental Science. Part IV: Sensor-Derived Phenology*, chap Near-Surface Sensor-Derived Phenology, pp 413–430. In: Schwartz (2013c), DOI 10.1007/978-94-007-6925-0_22
- Richter R, Schläpfer D (2016) *ATCOR-2/3 User Guide*. Version 9.1.0, Available online: https://www.rese-apps.com/pdf/atcor3_manual.pdf (accessed on 14 06 2017)
- Robel J, Graumann A (2014) NOAA KLM USER’S GUIDE with NOAA-N, N Prime and MetOp SUPPLEMENTS. Tech. rep., National Oceanic and Atmospheric Administration; National Environmental Satellite, Data and Information Service; National Climatic Data Center; Remote Sensing and Applications Division, Asheville, NC, USA
- Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421(6918):57–60, DOI 10.1038/nature01333
- Root TL, MacMynowski DP, Mastrandrea MD, Schneider SH (2005) Human-modified temperatures induce species changes: Joint attribution. *Proceedings of the National Academy of Sciences of the United States of America* 102(21):7465–7469, DOI 10.1073/pnas.0502286102
- Rossini M, Meroni M, Migliavacca M, Manca G, Cogliati S, Busetto L, Picchi V, Cescatti A, Seufert G, Colombo R (2010) High resolution field spectroscopy measurements for estimating gross ecosystem production in a rice field. *Agricultural and Forest Meteorology* 150(9):1283–1296, DOI 10.1016/j.agrformet.2010.05.011
- Rossini M, Cogliati S, Meroni M, Migliavacca M, Galvagno M, Busetto L, Cremonese E, Julitta T, Siniscalco C, Morra di Cella U, Colombo R (2012) Remote sensing-based estimation of gross primary production in a subalpine grassland. *Biogeosciences* 9(7):2565–2584, DOI 10.5194/bg-9-2565-2012
- Ryu Y, Verfaillie J, Macfarlane C, Kobayashi H, Sonnentag O, Vargas R, Ma S, Baldocchi DD (2012) Continuous observation of tree leaf area index at ecosystem scale using

- upward-pointing digital cameras. *Remote Sensing of Environment* 126:116–125, DOI 10.1016/j.rse.2012.08.027
- Saino N, Ambrosini R, Rubolini D, von Hardenberg J, Provenzale A, Hüppop K, Hüppop O, Lehikoinen A, Lehikoinen E, Rainio K, Romano M, Sokolov L (2011) Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings of the Royal Society B-Biological Sciences* 278(1707):835–842, DOI 10.1098/rspb.2010.1778
- Sakowska K, Vescovo L, Marcolla B, Juszczak R, Olejnik J, Gianelle D (2014) Monitoring of carbon dioxide fluxes in a subalpine grassland ecosystem of the Italian Alps using a multispectral sensor. *Biogeosciences* 11(17):4695–4712, DOI 10.5194/bg-11-4695-2014
- Sakowska K, Gianelle D, Zaldei A, MacArthur A, Carotenuto F, Miglietta F, Zampedri R, Cavagna M, Vescovo L (2015) WhiteRef: A New Tower-Based Hyperspectral System for Continuous Reflectance Measurements. *Sensors* 15(1):1088–1105, DOI 10.3390/s150101088
- Sakowska K, Juszczak R, Gianelle D (2016) Remote Sensing of Grassland Biophysical Parameters in the Context of the Sentinel-2 Satellite Mission. *Journal of Sensors* 2016(1):1–16, DOI 10.1155/2016/4612809
- Saxe H, Cannell M, Johnsen B, Ryan M, Vourlitis G (2001) Tree and forest functioning in response to global warming. *New Phytologist* 149(3):369–399, DOI 10.1046/j.1469-8137.2001.00057.x
- Schaber J (2002) Phenology in Germany in the 20th Century: Methods, Analyses and Models. PhD thesis, University of Potsdam, Potsdam, Germany, available online: <https://publishup.uni-potsdam.de/frontdoor/index/index/docId/49> (accessed on 19 September 2017)
- Schaber J (2012) *pheno: Auxiliary functions for phenological data analysis*. R package version 1.6, <http://CRAN.R-project.org/package=pheno> (accessed 09 January 2013)
- Schaber J, Badeck FW (2002) Evaluation of methods for the combination of phenological time series and outlier detection. *Tree Physiology* 22:973–982, DOI 10.1093/treephys/22.14.973

Bibliography

- Schaber J, Badeck FW (2003) Physiology-based phenology models for forest tree species in germany. *International Journal of Biometeorology* 47:193–201, DOI 10.1007/s00484-003-0171-5
- Schaber J, Badeck FW (2005) Plant phenology in germany over the 20th century. *Regional Environmental Change* 5:37–46, DOI 10.1007/s10113-004-0094-7
- Schnelle F (1955) *Pflanzen-Phänologie*, 1st edn. Probleme der Bioklimatologie, Akademische Verlagsgesellschaft Geest & Portig K.-G, Leipzig
- Schwartz MD (ed) (2013a) *Phenology: An Integrative Environmental Science. Part I: Phenological Data, Networks, and Research*. Springer Netherlands, Dordrecht, DOI 10.1007/978-94-007-6925-0
- Schwartz MD (ed) (2013b) *Phenology: An Integrative Environmental Science. Part III: Phenological Models and Techniques*. Springer Netherlands, Dordrecht, DOI 10.1007/978-94-007-6925-0
- Schwartz MD (ed) (2013c) *Phenology: An Integrative Environmental Science. Part IV: Sensor-Derived Phenology*. Springer Netherlands, Dordrecht, DOI 10.1007/978-94-007-6925-0
- Schwartz MD (ed) (2013d) *Phenology: An Integrative Environmental Science. Part VI: Applications of Phenology*. Springer Netherlands, Dordrecht, DOI 10.1007/978-94-007-6925-0
- Schwartz MD, Crawford TM (2001) Detecting energy balance modifications at the onset of spring. *Physical Geography* 22(5):394–409, DOI 10.1080/02723646.2001.10642751
- Schwartz MD, Ahas R, Aasa A (2006) Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology* 12(2):343–351, DOI 10.1111/j.1365-2486.2005.01097.x
- Sellers PJ, Tucker CJ, Collatz GJ, Los SO, Justice CO, Dazlich DA, Randall DA (1994) A Global 1-Degrees-by-1-Degrees NDVI Data Set for Climate Studies .2. The Generation of Global Fields of Terrestrial Biophysical Parameters from the NDVI. *International Journal of Remote Sensing* 15(17):3519–3545, DOI 10.1080/01431169408954343

- Shen M, Tang Y, Chen J, Zhu X, Zheng Y (2011) Influences of temperature and precipitation before the growing season on spring phenology in grasslands of the central and eastern Qinghai-Tibetan Plateau. *Agricultural and Forest Meteorology* 151(12):1711–1722, DOI {10.1016/j.agrformet.2011.07.003}
- Sherry R, Zhou X, Gu S, Arnone J, Schimel D, Verburg P, Wallace L, Luo Y (2007) Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences of the United States of America* 104(1):198–202, DOI 10.1073/pnas.0605642104
- Singer MC, Parmesan C (2010) Phenological asynchrony between herbivorous insects and their hosts: signal of climate change or pre-existing adaptive strategy? *Philosophical Transactions of the Royal Society B-Biological Sciences* 365(1555):3161–3176, DOI 10.1098/rstb.2010.0144
- Soudani K, le Maire G, Dufrêne E, François C, Delpierre N, Ulrich E, Cecchini S (2008) Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sensing of Environment* 112(5):2643–2655, DOI 10.1016/j.rse.2007.12.004, Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue
- Soudani K, Hmimina G, Delpierre N, Pontailier JY, Aubinet M, Bonal D, Caquet B, de Grandcourt A, Burban B, Flechard C, Guyon D, Granier A, Gross P, Heinesh B, Longdoz B, Loustau D, Moureaux C, Ourcival JM, Rambal S, Saint André L, Dufrêne E (2012) Ground-based Network of NDVI measurements for tracking temporal dynamics of canopy structure and vegetation phenology in different biomes. *Remote Sensing of Environment* 123(Supplement C):234–245, DOI 10.1016/j.rse.2012.03.012
- Spekat A, Enke W, Kreienkamp F (2007) Neuentwicklung von regional hoch aufgelösten Wetterlagen für Deutschland und Bereitstellung regionaler Klimaszenarios auf der Basis von globalen Klimasimulationen mit dem Regionalisierungsmodell WETTREG auf der Basis von globalen Klimasimulationen mit ECHAM5/MPI-OM T63L31 2010 bis 2100 für die SRES-Szenarios B1, A1B und A2. Endbericht, Umweltbundesamt, Förderkennzeichen 204 41 138
- Stanners DA, Bourdeau P (1995) Europe's environment: the Dobris assessment. European Environment Agency, Office for the Official Publications of the European

- Communities, Luxembourg, LU, available online: <https://www.eea.europa.eu/publications/92-826-5409-5/> (accessed on 17 Oct 2017)
- Stellmes M, Röder A, Udelhoven T, Hill J (2013) Mapping syndromes of land change in Spain with remote sensing time series, demographic and climatic data. *Land Use Policy* 30(1):685–702, DOI 10.1016/j.landusepol.2012.05.007
- Stöckli R, Vidale PL (2004) European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. *International Journal of Remote Sensing* 25(17):3303–3330, DOI 10.1080/01431160310001618149
- Tarpley JD (1991) The NOAA Global Vegetation Index product - A review. *Paleogeography, Paleoclimatology, Paleoecology (Global and Planetary Change)* 4(1):189–194, DOI 10.1016/0921-8181(91)90091-A
- Thackeray SJ, Sparks TH, Frederiksen M, Burthe S, Bacon PJ, Bell JR, Botham MS, Brereton TM, Bright PW, Carvalho L, Clutton-Brock T, Dawson A, Edwards M, Elliott JM, Harrington R, Johns D, Jones ID, Jones JT, Leech DI, Roy DB, Scott WA, Smith M, Smithers RJ, Winfield IJ, Wanless S (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16(12):3304–3313, DOI 10.1111/j.1365-2486.2010.02165.x
- Thayn JB (2012) Assessing vegetation cover on the date of satellite-derived start of spring. *Remote Sensing Letters* 3(8):721–728, DOI 10.1080/2150704X.2012.674227
- Tortini R, Hilker T, Coops N, Nesic Z (2015) Technological Advancement in Tower-Based Canopy Reflectance Monitoring: The AMSPEC-III System. *Sensors* 15(12):32,020–32,030, DOI 10.3390/s151229906
- Tucker CJ (1979) Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment* 8:127–150, DOI 10.1016/0034-4257(79)90013-0
- USGS (2018) LANDSAT Web Page. U.S. Geological Survey, Department of the Interior, Baltimore, MD, USA; Available online: <https://landsat.usgs.gov/> (accessed on 23 Feb 2018)
- Vanbeverem SP, Bloemen J, Balzarolo M, Broeckx LS, Sarzi-Falchi I, Verlinden MS, Ceulemans R (2016) A comparative study of four approaches to assess phenology of *Populus* in a short-rotation coppice culture. *iForest* 9(5):682–689, DOI 10.3832/ifor1800-009

- Vermote EF, Kotchenova SY (2008a) Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research* 113(D23S90):1–12, DOI 10.1029/2007JD009662
- Vermote EF, Kotchenova SY (2008b) *MOD09 (Surface Reflectance) User's Guide, 1.1 ed.* MODIS Land Surface Reflectance Science Computing Facility: Greenbelt, MD, USA, 2008
- Vermote, E and Wolfe, R (2015a) MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V005. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (<https://lpdaac.usgs.gov>), available online: <https://e4ftl01.cr.usgs.gov/MOLT/> (accessed on 16 January 2017)
- Vermote, E and Wolfe, R (2015b) MYD09GQ MODIS/Aqua Surface Reflectance Daily L2G Global 250m SIN Grid V005. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (<https://lpdaac.usgs.gov>), available online: <https://e4ftl01.cr.usgs.gov/MOLA/> (accessed on 16 January 2017)
- Viovy N, Arino O, Belward AS (1992) The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing* 13(8):1585–1590, DOI 10.1080/01431169208904212
- Visser ME, Both C (2005) Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B-Biological Sciences* 272(1581):2561–2569, DOI 10.1098/rspb.2005.3356
- Visser ME, Holleman LJM (2001) Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268(1464):289–294, DOI 10.1098/rspb.2000.1363
- Vitasse Y, François C, Delpierre N, Dufrêne E, Kremer A, Chuine I, Delzon S (2011) Assessing the effects of climate change on the phenology of European temperate trees. *Agricultural and Forest Meteorology* 151(7):969–980, DOI 10.1016/j.agrformet.2011.03.003
- Walther GR (2004) Plants in a warmer world. *Perspectives in plant ecology evolution and systematics* 6(3):169–185, DOI 10.1078/1433-8319-00076

Bibliography

- Wang JY (1960) A critique of the heat unit approach to plant response studies. *Ecology* 41:785–790, DOI 10.2307/1931815
- Wang X, Piao S, Xu X, Ciais P, MacBean N, Myneni RB, Li L (2015) Has the advancing onset of spring vegetation green-up slowed down or changed abruptly over the last three decades? *Global Ecology and Biogeography* 24(6):621–631, DOI {10.1111/geb.12289}
- Wareing PF (1956) Photoperiodism in woody plants. *Annual Review of Plant Physiology* 7:191–214, DOI 10.1146/annurev.pp.07.060156.001203
- Wareing PF, Saunders PF (1971) Hormones and dormancy. *Annual Review of Plant Physiology* 22:261–288, DOI 10.1146/annurev.pp.22.060171.001401
- White K, Pontius J, Schaberg P (2014) Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. *Remote Sensing of Environment* 148:97–107, DOI {10.1016/j.rse.2014.03.017}
- White MA, Nemani RR (2006) Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment* 104(1):43–49, DOI 10.1016/j.rse.2006.04.014
- White MA, Thornton PE, Running SW (1997) A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* 11(2):217–234, DOI 10.1029/97GB00330
- White MA, De Beurs KM, Didan K, Inouye DW, Richardson AD, Jensen OP, O’Keefe J, Zhang G, Nemani RR, Van Leeuwen WJD, Brown JF, De Wit A, Schaepman M, Lin X, Dettinger M, Bailey AS, Kimball J, Schwartz MD, Baldocchi DD, Lee JT, Lauenroth WK (2009) Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology* 15(10):2335–2359, DOI 10.1111/j.1365-2486.2009.01910.x
- Wingate L, Ogée J, Cremonese E, Filippa G, Mizunuma T, Migliavacca M, Moisy C, Wilkinson M, Moureaux C, Wohlfahrt G, Hammerle A, Hörtnagl L, Gimeno C, Porcar-Castell A, Galvagno M, Nakaji T, Morison J, Kolle O, Knohl A, Kutsch W, Kolari P, Nikinmaa E, Ibrom A, Gielen B, Eugster W, Balzarolo M, Papale D, Klumpp K, Köstner B, Grünwald T, Joffre R, Ourcival JM, Hellstrom M, Lindroth A, Charles G, Longdoz B, Genty

- B, Levula J, Heinesch B, Sprintsin M, Yakir D, Manise T, Guyon D, Ahrends H, Plaza-Aguilar A, Guan J, Grace J (2015) Interpreting canopy development and physiology using the EUROPhen camera network at flux sites. *Biogeosciences Discuss* 12(10):7979–8034, DOI 10.5194/bgd-12-7979-2015
- Wittich KP, Kraft M (2008) The normalised difference vegetation index obtained from agrometeorological standard radiation sensors: a comparison with ground-based multi-band spectroradiometer measurements during the phenological development of an oat canopy. *International Journal of Biometeorology* 52(3):167–177, DOI {10.1007/s00484-007-0108-5}
- Wollschläger U, Attinger S, Borchardt D, Brauns M, Cuntz M, Dietrich P, Fleckenstein JH, Friesen K, Friesen J, Harpke A, Hildebrandt A, Jäckel G, Kamjunke N, Knöller K, Kögler S, Kolditz O, Krieg R, Kumar R, Lausch A, Liess M, Marx A, Merz R, Mueller C, Musolff A, Norf H, Oswald S, Rebmann C, Reinstorf F, Rode M, Rink K, Rinke K, Samaniego L, Vieweg M, Vogel HJ, Weitere M, Werban U, Zink M, Zacharias S (2016) The Bode hydrological observatory: a platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory. *Environmental Earth Sciences* 76(1):29, DOI 10.1007/s12665-016-6327-5
- Xu X, Conrad C, Doktor D (2017) Optimising Phenological Metrics Extraction for Different Crop Types in Germany Using the Moderate Resolution Imaging Spectrometer (MODIS). *Remote Sensing* 9(3), DOI 10.3390/rs9030254
- Yang H, Yang X, Zhang Y, Heskell MA, Lu X, Munger JW, Sun S, Tang J (2017) Chlorophyll fluorescence tracks seasonal variations of photosynthesis from leaf to canopy in a temperate forest. *Global Change Biology* 23(7):2874–2886, DOI 10.1111/gcb.13590
- Yang X, Tang J, Mustard JF, Lee JE, Rossini M, Joiner J, Munger JW, Kornfeld A, Richardson AD (2015) Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest: Fluorescence and photosynthesis. *Geophysical Research Letters* 42(8):2977–2987, DOI 10.1002/2015GL063201
- Zhang G, Zhang Y, Dong J, Xiao X (2013) Green-up dates in the Tibetan Plateau have continuously advanced from 1982 to 2011. *Proceedings of the National Academy of Sciences* 110(11):4309–4314, DOI 10.1073/pnas.1210423110

Bibliography

- Zhang Q, Xiao X, Braswell B, Linder E, Ollinger S, Smith ML, Jenkins JP, Baret F, Richardson AD, Moore III B, Minocha R (2006) Characterization of seasonal variation of forest canopy in a temperate deciduous broadleaf forest, using daily MODIS data. *Remote Sensing of Environment* 105(3):189–203, DOI 10.1016/j.rse.2006.06.013
- Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JCF, Gao F, Reed BC, Huete A (2003) Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment* 84(3):471–475, DOI 10.1016/S0034-4257(02)00135-9
- Zhang X, Wang J, Gao F, Liu Y, Schaaf C, Friedl M, Yu Y, Jayavelu S, Gray J, Liu L, Yan D, Henebry GM (2017) Exploration of scaling effects on coarse resolution land surface phenology. *Remote Sensing of Environment* 190:318–330, DOI 10.1016/j.rse.2017.01.001
- Zhang X, Friedl MA, Schaaf CB (2006) Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research-Biogeosciences* 111(G4), DOI {10.1029/2006JG000217}
- Zhou L, Tucker CJ, Kaufmann RK, Slayback D, Shabanov NV, Myneni RB (2001) Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research: Atmospheres* 106(D17):20,069–20,083, DOI 10.1029/2000JD000115

List of Figures

1.1	MODIS and Sentinel-2 NDVI image of the area around Hohes Holz and Grosser Bruch	7
1.2	Study site locations in Germany	10
1.3	Europe map of Köppen climate classification	11
1.4	Study site locations and land cover within Sentinel-2 tile T32UPC	13
2.1	Phenology observation stations in Germany	23
2.2	Statistics of phenological observations from 1951-2009	24
2.3	PIM compound development	26
2.4	Combining phenological time series	29
2.5	PIM parameter sensitivity	31
2.6	PIM parameter correlation	33
2.7	PIM predictions versus observations for beech and birch	34
2.8	PIM predictions versus observations for chestnut and oak	35
2.9	Modelled phenological trends for period 2002-2100	37
2.10	Relationship of BB development and temperature sums	38
3.1	Phenological observation stations (grey triangles) within Germany which were active between the years 1951 to 2009. For each natural region (black lines, a region which is considered to be homogeneous with respect to topography, soil, climate and vegetation) at least one (or more) combined time series was computed (red triangles) and subsequently used for phenological modelling.	53

3.2	Implemented seasonal evolution of the PROSAIL parameters C_{ab} , C_{ar} , C_{brown} , C_w , C_m and N . PROSAIL simulations were performed for four different land-cover types: winter crops (black line), summer crops (red line), pastures (green line) and forests (blue line).	55
3.3	Simulated seasonal NDVI profiles of an exemplary year considering homogeneous land-cover including winter crops (purple line), summer crops (orange line), pastures (green line) and deciduous forest (darkgreen line) as well as the four cases of heterogeneous land-cover <i>Equal area</i> (solid black), <i>Arbitrary area</i> (long dashed), <i>Uneven area</i> (short dashed) and <i>Spatial trend</i> (short/long dashed).	58
3.4	Trend differences (simulated - extracted) [days/year] across methods for i) a single land-cover within a satellite pixel (Dec. forest, Summer crop, Pastures, Winter crop) and ii) mixed land-cover within a satellite pixel (<i>Spatial trend</i> , <i>Uneven area</i> , <i>Equal area</i> , <i>Arbitrary area</i>). Trend estimates are based on selected simulated NDVI values (for 19 years, 100 repetitions) – according to the likelihood of clear sky observations – with an induced trend of 1 day/year towards earlier green-up.	62
3.5	Average goodness-of-fit, R^2 , for CORINE land-use classes (x-axis) and approximating methods (<i>Lg</i> , <i>Sg</i> , <i>FFT</i> , <i>Gs</i> , <i>SG</i>), i.e. excluding <i>Li</i> , <i>Sp</i> . R^2 values are based on the period 1989-2007 for fits on daily NDVI observations. The dashed blue line indicates R^2 values of 0.6.	63
3.6	Mean d_{GU} of the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold (<i>kl65</i>) using linear interpolation (<i>Li</i>).	64
3.7	Mean d_{GU} of the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold (<i>kl65</i>) for all other methods. All plots show the difference of the respective method to <i>Li</i> (i.e. Figure 3.6).	65
3.8	Trend [days/year] of all tested methods for the period 1989-2007 based on daily NDVI observations for areas of deciduous forest in Germany (light grey) and Europe (dark grey) using a dynamic (local, <i>kl65</i>) threshold. Trend estimates for ground observations are shown in green and for modelled phenology (both Germany) in red boxplots. The dashed blue line divides positive and negative trend estimates.	66

3.9	Trend [days/year] of the period 1989-2007 for Europe and the Mediterranean based on daily NDVI observations and a local threshold ($kl65$) for all methods: Li , Sp , FFT , SG , Gs , Sg and Lg	68
3.10	d_{GU} trend (days/year) over the period 1989-2007 for Central/Western Europe and the Mediterranean based on daily NDVI observations and a local threshold ($kl65$) across major CORINE land-use classes.	70
4.1	Pictures of sensor system mounted at the eddy flux tower at 'Hohes Holz' .	86
4.2	Schematic overview of sensor setup for unattended hyperspectral and multispectral measurements.	87
4.3	Sentinel-2A NDVI time series from 2016 with 10-m resolution processed with ATCOR 2/3.	90
4.4	NDVI time series of 2015 and 2016 derived from different sensor systems at 'Hohes Holz'	92
4.5	Pictures of the area around the eddy flux tower 'Hohes Holz'.	93
4.6	Scatter-plot of NDVI derived from the hyperspectral sensor system and NDVI from the multispectral sensor system	94
4.7	Scatter-plot of NDVI derived from the hyperspectral sensor system and satellite NDVI products	95
4.8	Pictures of phenological state at 'Hohes Holz'	96
5.1	Overview of the different sensors, calibration and assessment steps	111
5.2	Overview of observations of different systems under same light conditions .	115
5.3	Daily ratios between multi- and corresponding resampled hyperspectral observations	117
5.4	Hyper- vs. multispectral reflectance	119
5.5	Validation of observations with portable field spectrometer measurements .	122
5.6	Comparison of reference hyperspectral solar irradiance with calibration result of the automated systems at the two sites	133
5.7	Example of white reference measurement containing a weak vegetation signal contribution before and after correction	134
5.8	Comparison of multispectral solar irradiance between the grassland and forest sites on selected sunny days	135
5.9	Comparison of hyper- and multispectral downwelling irradiance observations before and after correction	136

List of Figures

5.10 Comparison of hyper- and multispectral upwelling irradiance observations before and after correction	136
5.11 Illustration of discontinuity correction of hyperspectral downwelling irradi- ance observations at the grassland site	137
5.12 Illustration of grassland reflectance results after discontinuity correction . .	137
5.13 Illustration of the white reference observations	138

List of Tables

2.1	Characterisation of temperature datasets	22
2.2	PIM parameter sets	27
2.3	Errors and r^2 of PIM fits	28
2.4	Correlations of photo-period, temperature and chilling days with bud burst .	30
2.5	Modelled phenological trends	36
3.1	The four land-cover mixture type cases for which NDVI values were simulated with PROSAIL: Equal area, uneven area, arbitrary area and spatial trend.	56
3.2	Methods used for extracting phenological metrics from NDVI observations. All methods are implemented in 'R' either using R-functions or embedded C-code and are freely available through the R-package 'phenex'. An initial dynamic filtering (BISE) is applied on all satellite based daily NDVI data beforehand.	59
3.3	Mean absolute deviation of computed green-up dates (d_{GU}) across methods (rows) and thresholds (columns) for broadleaf forests in Germany compared to respective ground observations (d_{BB}).	60
3.4	Trends (days/year) for simulated NDVI time series with daily resolution excluding sensor noise.	61
3.5	Percentage of pixels per method falling into discrete trend categories (d/y, days per year) also used for mapping (s. Figure 3.9).	69
4.1	Band configuration of the multispectral sensor.	85
4.2	Squared Pearson correlation coefficients of NDVI product combinations from different sensors for the forest site 'Hohes Holz'.	94
4.3	Overview of extracted phenological metrics	97

List of Tables

5.1	Specifications of the multispectral channels.	110
5.2	Comparison of corrected and uncorrected observations	120
5.3	Evaluation of grassland reflectance results after discontinuity correction . .	132
5.4	Overview of continuous correction results for reflectance using different single or multiple channels in the correction procedure	132

List of Abbreviations and Variables

Abbreviations

ASD	Analytical Spectral Devices Inc., Boulder, CO, USA
ATCOR	Atmospheric & Topographic Correction, RESE®
AVHRR	Advanced Very High Resolution Radiometer
BB	Bud burst
BISE	Best Index Slope Extraction (Viovy et al., 1992)
C_{ab}	Chlorophyll content (biophysical parameter, see Section 3.2.4)
C_{ar}	Carotenoid content (biophysical parameter, see Section 3.2.4)
C_{brown}	Brown pigment content (biophysical parameter, see Section 3.2.4)
CET	Central European Time
C_m	Dry matter content (biophysical parameter, see Section 3.2.4)
CORINE	Coordinated Information on the Environment (European Environment Agency, 2013; Bundesamt für Kartographie und Geodäsie, 2012)
CV	Coefficient of variation
CW	Central wavelength

List of Abbreviations and Variables

C_w	Equivalent water thickness (biophysical parameter, see Section 3.2.4)
DDSRR	Direct to diffuse solar radiation ratio
DFOV	Dual-field-of-view
DOY	Julian day (Day Of the Year)
DWD	German Weather Service (Deutscher Wetterdienst)
EC	Eddy covariance
ESA	European Space Agency
EVI	Enhanced Vegetation Index
FFT	Fast-Fourier Transformation method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
FOV	Field of view
FWHM	Full-width at half-maximum
GIMMS	Global Inventory Monitoring and Modelling System (Pinzon and Tucker, 2014)
GloVis	USGS Global Visualization Viewer (https://glovis.usgs.gov/)
GPP	Gross Primary Productivity
Gs	Gaussian function method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
GT	Global threshold method (depiction of phenological metrics, see Section 3.2.5)
GU	Green-up date
GVI	Global Vegetation Index
ICOS	Integrated Carbon Observation System (ICOS ERIC, 2017)

IPCC	Intergovernmental Panel on Climate Change
kg35, kg40	Phenological metrics depicted with Global Threshold method with threshold value of 0.35 or 0.4, respectively (see Section 3.2.5)
kl55, kl60,kl65	Phenological metrics depicted with Local Threshold method with threshold value of 0.55, 0.6 or 0.65, respectively (see Section 3.2.5)
LAI	Leaf Area Index
LC	Land-cover type
Li	Linear Interpolation method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
Lg	Logistic function method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
LSP	Land surface phenology
LT	Local threshold method (depiction of phenological metrics, see Section 3.2.5)
MAE	Mean absolute error
MEDOKADS	Mediterranean Extended Daily One Km AVHRR Data Set (see Section 3.2.1)
MOD09GQ	MODIS Terra daily surface reflectance dataset (Vermote, E. and Wolfe, R., 2015a)
MODIS	Moderate Resolution Imaging Spectroradiometer
MVC	Maximum Value Composite (Holben, 1986)
MYD09GQ	MODIS Aqua daily surface reflectance dataset (Vermote, E. and Wolfe, R., 2015b)
NASA	National Aeronautics and Space Administration
N	Leaf mesophyll structure parameter of PROSAIL (Jacquemoud (1993), see Section 3.2.4)

List of Abbreviations and Variables

NDVI	Normalized Difference Vegetation Index
NIR	Near-infrared
NOAA	National Oceanic and Atmospheric Administration
obs	Observation
PAR	Photosynthetically Active Radiation
PIM	Promoter-Inhibitor-Model (Schaber and Badeck, 2003)
PPODB	Plant Phenological Online Database (Dierenbach et al., 2013)
PRI	Photochemical Reflectance Index
PROSAIL	A radiative transfer model built through coupling the leaf optical properties model PROSPECT and the canopy bidirectional reflectance model SAIL (Jacquemoud et al., 2009)
r²	Squared Pearson correlation coefficient (Pearson, 1920)
REMO-UBA	Regional climate simulation model of the Max Planck Institute for Meteorology (Hamburg, Germany) in cooperation with the Umweltbundesamt (Dessau-Roßlau, Germany)
RMSE	Root mean squared error
rRMSE	Root mean squared error relative to mean value
SAR	Sum of absolute residuals
sd	Standard deviation
SEN	Senescence date
Sen2cor	Processor for Sentinel-2 Level 2A product generation and formatting

Sg	Sigmoidal function method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
SG	Savitzky-Golay filtering method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
SIF	Sun-Induced chlorophyll Fluorescence
Sp	Monotone Hermite Spline Interpolation method for the extraction of phenological metrics (see Section 3.2.5 and table 3.2)
SPOT	Système Probatoire d’Observation de la Terre
SWDR	Shortwave Downward Radiation
TERENO	Terrestrial Environmental Observatories
UFZ	Helmholtz-Centre for Environmental Research - UFZ (UmweltForschungsZentrum)
USB	Universal Serial Bus
USGS	United States Geological Survey
UTC	Universal Time, Coordinated
VCF	Vegetation Cover Fraction (Thayn, 2012)
WETTREG	Regional climate simulation model (WETTerlagen-basierte REGIONalisierung), described in Spekat et al. (2007)
WMO	World Meteorological Organization
WR	White reference

Variables

a_i with $i = 1, \dots, 4$	Scaling variables of the phenological model PIM for the forcing and breakdown terms of inhibiting and promoting compound (see Section 2.2.3)
---------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------

List of Abbreviations and Variables

β_0	Intercept of the linear regression model for trend analysis in Section 3.2.6
β_1	Slope of the linear regression model for trend analysis in Section 3.2.6
BB_{mod}	Modelled bud burst date (as Julian day, see Chapter 2)
BB_{NA}	Error value of the phenological model PIM (see Section 2.2.3)
BB_{obs}	Observed bud burst date (as Julian day, see Chapter 2)
c	Slope of the first inflection point of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
d	Slope of the second inflection point of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
d_{BB}	Record of bud burst day (see Section 3.1)
d_{GU}	Satellite derived green-up dates (see Section 3.1)
$data_{cts}$	Combined phenological time series (see Section 2.2.2)
$data_{fit}$	Phenological time series used to train the phenological model PIM (see Section 2.2.2)
$data_{val}$	Phenological time series used to validate the results of the phenological model PIM (see Section 2.2.2)
ε	Error variable of the linear regression model for trend analysis in Section 3.2.6
F	Fitness value and the respective objective function used to fit the phenological model PIM (see Section 2.2.4)
f_{chill}	Weighted chilling days (see Section 2.2.3 and Equation 2.5)
$f(T)$	Function of temperature dependency in the phenological model PIM (see Section 2.2.3)
$g(L)$	Function of photo-period dependency in the phenological model PIM (see Section 2.2.3)
I	Value of the growth inhibiting compound used in the phenological model PIM (see Section 2.2.3)

Irr_i	Downwelling irradiance measurements of sensor channel i selected by the methodology described in Section 4.2.2
$Irr_{i,change}$	Absolute irradiance change of sensor channel i (see Section 4.2.2)
$\overline{Irr_i}$	Mean irradiance of sensor channel i (see Section 4.2.2)
j	Number of model parameters used in the phenological model PIM (see Section 2.2.4)
l_k	Lower bound of the parameter x_k of the phenological model PIM (see Section 2.2.5)
k	1) Identifier (Number) of the parameter x_k in the phenological model PIM (see Section 2.2.5) 2) asymptotic value of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4 3) Threshold value in Equation 3.3 in Section 3.2.5
L	Day length (photo-period) in hours used in the phenological model PIM (see Section 2.2.3)
$NDVI_{aqua}$	NDVI data derived from MODIS aqua product MYD09GQ (see Section 4.2.3)
$NDVI_{aqua}^{bise}$	NDVI values selected by the Best Index Slope Extraction (Viovy et al., 1992) from $NDVI_{aqua}$ (see Section 4.2.4)
$NDVI_{hyper}$	NDVI data from the hyperspectral sensor system generated by the methodology described in Section 4.2.2
$NDVI_{multi}$	NDVI data from the multispectral sensor system generated by the methodology described in Section 4.2.2
$NDVI_{terra}$	NDVI data derived from MODIS terra product MOD09GQ (see Section 4.2.3)
$NDVI_{terra}^{bise}$	NDVI values selected by the Best Index Slope Extraction (Viovy et al., 1992) from $NDVI_{terra}$ (see Section 4.2.4)
$NDVI_{s2a,10}^{atcor}$	NDVI data derived from Sentinel-2a Level 2a data with spatial resolution of 10 m, atmospherically-, terrain- and cirrus-corrected with RESE©ATCOR 2/3 (see Section 4.2.3)
$NDVI_{s2a,20}^{atcor}$	NDVI data derived from Sentinel-2a Level 2a data with spatial resolution of 20 m, atmospherically-, terrain- and cirrus-corrected with RESE©ATCOR 2/3 (see Section 4.2.3)

List of Abbreviations and Variables

$NDVI_{min}$	Minimum NDVI (see Section 3.2.5)
$NDVI_{max}$	Maximum NDVI (see Section 3.2.5)
p	Date of the first inflection point of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
P	Value of the growth promoting compound used in the phenological model PIM (see Section 2.2.3)
q	Date of the second inflection point of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
r^2	Squared Pearson correlation coefficient (Pearson, 1920)
Rad_i	Upwelling radiance measurements of sensor channel i selected by the methodology described in Section 4.2.2
$Refl_i$	Reflectance factor of sensor channel i (see Section 4.2.2 and Equation 4.1)
S_k	Parameter sensitivity for each parameter x_k of the phenological model PIM (see Section 2.2.5 and Equation 2.7)
t	Time value (as Julian day) of the phenological model PIM (see Section 2.2.3) and the NDVI simulation in Equation 3.1 in Section 3.2.4
t_0	Starting date (as Julian day) of the phenological model PIM (see Section 2.2.3)
t_{end}	Fixed end date (as Julian day) for the calculation of T_{Σ} (see Section 2.3.4 and Equation 2.8)
T	Daily mean temperature
T_{Σ}	Temperature sum (see Section 2.3.4 and Equation 2.8)
$T_{min,I}$	Temperature threshold used in the phenological model PIM determining the minimum temperature for the breakdown of the inhibiting compound (see Section 2.2.3)
$T_{max,I}$	Temperature threshold used in the phenological model PIM determining the maximum temperature for the breakdown of the inhibiting compound (see Section 2.2.3)
$T_{opt,I}$	Temperature threshold used in the phenological model PIM determining the optimal temperature for the breakdown of the inhibiting compound (see Section 2.2.3)

$T_{min,P}$	Temperature threshold used in the phenological model PIM determining the minimum temperature for the forcing of the promoting compound (see Section 2.2.3)
$T_{max,P}$	Temperature threshold used in the phenological model PIM determining the maximum temperature for the forcing of the promoting compound (see Section 2.2.3)
$T_{opt,P}$	Temperature threshold used in the phenological model PIM determining the optimal temperature for the forcing of the promoting compound (see Section 2.2.3)
$thres_{inc}$	Parameter of the Best Index Slope Extraction (Viovy et al., 1992) determining the maximum NDVI increase allowed per day (see Section 4.2.3)
u_k	Upper bound of the parameter x_k of the phenological model PIM (see Section 2.2.5)
vb	The base NDVI value before green-up of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
ve	The base NDVI value at the end of senescence of the logistic function (Fischer, 1994) in Equation 3.1 in Section 3.2.4
x_k with $k = 1, \dots, j$	Parameter set of phenological model PIM (see Section 2.2.4)
X	Explanatory variable of the linear regression model for trend analysis in Section 3.2.6
Y	Dependent variable of the linear regression model for trend analysis in Section 3.2.6

Declaration of Authorship

I declare that this thesis and the work presented in it are my own and has been generated without inadmissible aid as the result of my own scientific research. I confirm that consulted or quoted work of others is always clearly cited and that all main sources of help are acknowledged. Where the thesis is based on work done by myself jointly with others, i clarified the contribution of each author. I declare that neither this thesis nor any part of it have previously been submitted for a degree or any other qualification at this University or any other institution, unless stated otherwise. I confirm that I did no doctorate in the past.

Eigenständigkeitserklärung

Ich versichere hiermit, dass die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche kenntlich gemacht wurden. Personen, von denen ich bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts Unterstützungsleistung erhalten habe, wurden je nach Beitrag als Koautoren oder in der Danksagung genannt. Ich versichere, dass keine weitere Personen, außer den bereits genannten, bei der geistigen Herstellung der vorliegenden Arbeit beteiligt waren, insbesondere nicht die Hilfe eines Promotionsberaters in Anspruch genommen wurde und weitere Personen von mir oder in meinem Auftrag weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Ich versichere, dass die vorgelegte Arbeit oder Teile daraus weder im Inland noch im Ausland in gleicher oder in ähnlicher Form einer anderen Prüfungsbehörde zum Zwecke einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Ich bestätige, dass bisher kein Promotionsversuch meiner Person stattgefunden hat.

