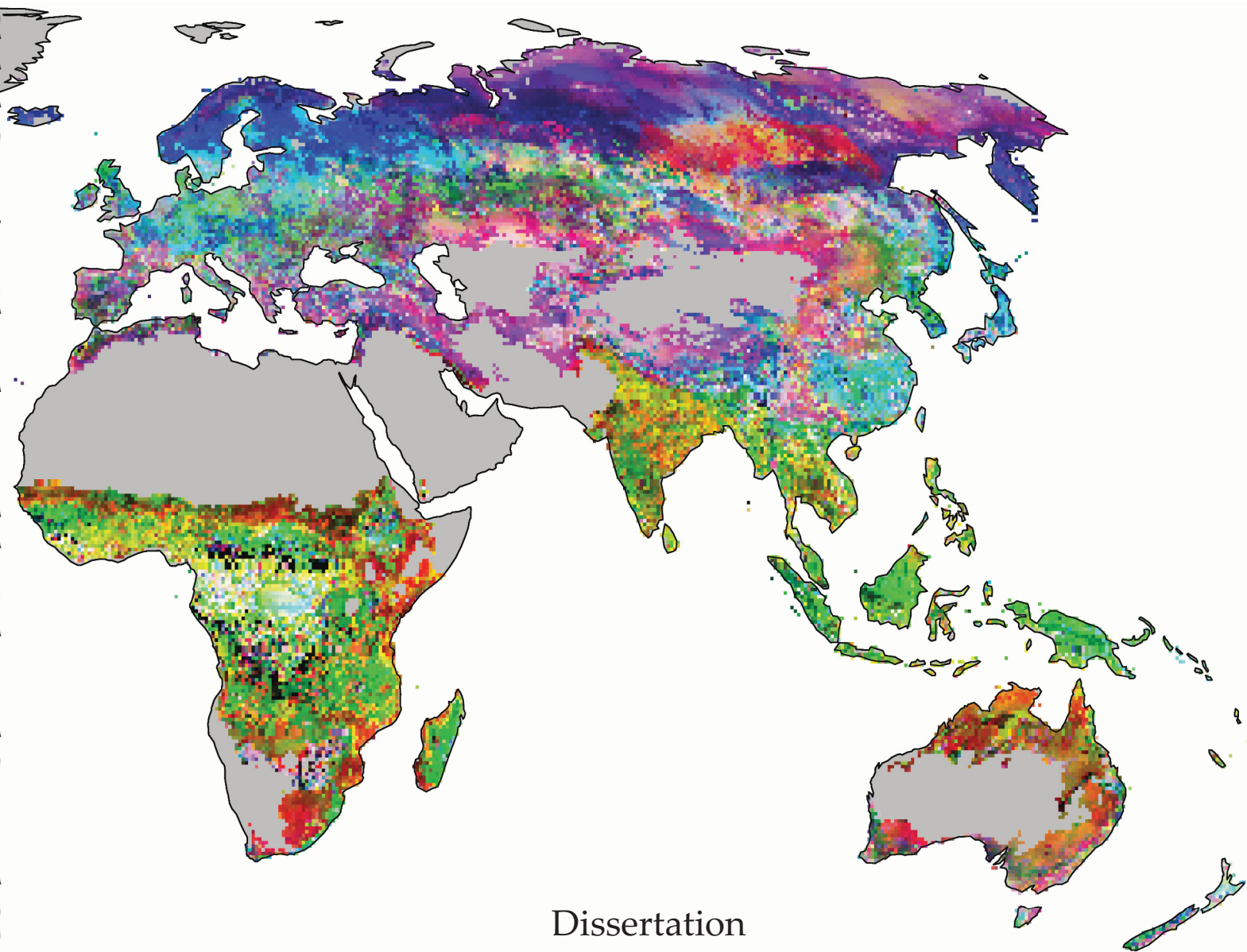


# Controls on Global Greening, Phenology and the Enhanced Seasonal CO<sub>2</sub> Amplitude

Integrating Decadal Satellite Observations  
and Global Ecosystem Models



Dissertation

Matthias Forkel

# Controls on Global Greening, Phenology and the Enhanced Seasonal CO<sub>2</sub> Amplitude

Integrating Decadal Satellite Observations  
and Global Ecosystem Models

Kumulative Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt dem Rat der Chemisch-Geowissenschaftlichen Fakultät

der Friedrich-Schiller-Universität Jena

von Matthias Forkel, M. Sc. (Geoinformatik)

geboren am 12. Juni 1985 in Zittau

#### Gutachter

1. Prof. Dr. Christiane Schmullius  
Institut für Geographie, Friedrich-Schiller-Universität Jena
2. Prof. Dr. Markus Reichstein  
Max-Planck-Institut für Biogeochemie Jena
3. Prof. Dr. Wolfgang Lucht  
Potsdam-Institut für Klimafolgenforschung und Humboldt-Universität Berlin

Tag der öffentlichen Verteidigung: 1. Juni 2015

In dankbarer Erinnerung an  
meinen Onkel und Wegweiser Rainer Erber (\* 1944 – † 2013) und  
meine „Tanti“ Anneliese Feurich (\* 1919 – † 2014).

In Liebe für Lenka und meine Tochter Helena (\* 2014).

*„Palaver mit zehn Fürsten und hundert Akademien, bis man irgendwo ein Barometer aufstellen dürfe. Das sei nicht Wissenschaft [schnaubte Gauß]. Ach, rief Humboldt, was sei Wissenschaft denn dann? Gauß sog an der Pfeife. Ein Mann allein am Schreibtisch.“*

(Kehlmann, 2007, "Die Vermessung der Welt", p.247)

# Contents

A full Table of Contents can be found at page 15.

<b>Contents</b> .....	5
<b>Summary</b> .....	6
<b>Zusammenfassung</b> .....	10
<b>Table of contents</b> .....	15
<b>Figures</b> .....	20
<b>Tables</b> .....	25
<b>List of manuscripts and publications</b> .....	26
<b>1 Introduction</b> .....	29
<b>2 Trend change detection in NDVI time series: effects of inter-annual variability and methodology</b> .....	63
<b>3 Identifying environmental controls on vegetation greenness phenology through model-data integration</b> .....	103
<b>4 Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness</b> .....	147
<b>5 Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere</b> .....	179
<b>6 Conclusions and future research</b> .....	191
<b>7 Supplement of chapter 3</b> .....	205
<b>8 Supporting information of chapter 4</b> .....	241
<b>9 Methods and extended data of chapter 5</b> .....	250
<b>References</b> .....	265
<b>Acknowledgement / Danksagung / Poděkování</b> .....	317
<b>Author contributions to the manuscripts</b> .....	318
<b>Curriculum vitae</b> .....	321
<b>Selbstständigkeitserklärung</b> .....	323

## Summary

**Background:** Satellite observations demonstrate wide-spread changes in terrestrial vegetation since the 1980s. For example, vegetation activity (or greenness) increases across the globe (“greening”). Observations also show regional decreases (“browning”) in parts of boreal forests and temperate grasslands, and changes in phenology, like an earlier beginning or a delayed end of the growing season in temperate and boreal forests. Satellite-derived indices of vegetation activity are indicative of changes in photosynthesis, and thus the terrestrial carbon cycle. For example, it has been suggested that greening and enhanced photosynthetic carbon uptake in high-latitude ecosystems is causing the observed increase in the seasonal amplitude of atmospheric CO<sub>2</sub>. However, the relationship between greening and the seasonality of atmospheric CO<sub>2</sub> has not yet been assessed while comprehensively accounting for other possible controls such as changes in agricultural productivity.

Moreover, the reliability of satellite-observed greening and browning trends has been contested, because regional trends cannot be detected in all satellite datasets and results might also differ depending on the chosen statistical analysis method. For example, browning of boreal forests in North America can be only detected in some datasets, and lead to different competing explanations for these trends. An overall greening and an earlier start of the growing season in high-latitude ecosystems have been explained by long-term increases in temperature. However, although global ecosystem models mostly rely on temperature relationships to simulate phenology, they do not reproduce observations of seasonal to long-term greenness dynamics well. This suggests that it is necessary to comprehensively consider several factors in global ecosystem models in order to represent and explain the observed seasonal to decadal changes in vegetation activity.

**Research aims:** The overall aim of this doctoral research is to better understand recent dynamics of land surface phenology and greenness in the Earth system by integrating satellite observations with global ecosystem models. Specifically, four research questions are addressed (Figure 1.2 and chapter 1.5). The first question aims to assess the contribution of observational uncertainties to the quantification of dynamics in land surface phenology and greenness. Observational uncertainties arise from

differences between several satellite datasets and from varying accuracies of time series analysis methods that are used to detect trends and trend changes, or to detect phenology events. The second aim is to improve the representation of phenology in a global ecosystem model (specifically LPJmL, Lund-Potsdam-Jena managed Lands), and to assess the impact of these model improvements on inter-annual and decadal changes in vegetation greenness and carbon cycle simulations. The third aim is to identify the regional and global importance of several controls on land surface phenology and greenness at multiple time scales (i.e. seasonality, inter-annual variability and trends). Finally, the fourth aim is to comprehensively assess the contribution of recent trends in vegetation greenness and productivity on the increasing amplitude of atmospheric CO<sub>2</sub>.

**Data, Methods and Models:** A variety of datasets and methods was used to address the research questions. Regarding the first aim, newly available satellite datasets of vegetation greenness were used together with a variety of time series analysis methods to assess observational uncertainties. For the second aim, a new phenology module (called LPJmL-GSI) was implemented within LPJmL and several observational data sets were integrated into the model to improve diagnostic model simulations. The new phenology module comprehensively accounts for effects of temperature, radiation and water availability on the seasonal development of canopies. Regarding the third aim, the improved model was applied in a factorial experiment to assess the regional and global importance of different controls on land surface phenology and greenness on multiple time scales. The considered controls are temperature, radiation, water availability, land use/land cover change, permafrost dynamics, fire, and CO<sub>2</sub> fertilization. Finally, LPJmL was used together with an atmospheric transport model (TM3) to comprehensively assess the role of several factors, including greening in boreal and arctic ecosystems, on trends in the seasonal amplitude of atmospheric CO<sub>2</sub>.

**Main results:** The thesis provides a more comprehensive assessment of 1) the controls on vegetation activity on seasonal to decadal time scales and 2) the consequent effects on the seasonal amplitude of atmospheric CO<sub>2</sub>.

It is essential to rigorously assess observational uncertainties for a robust detection of changes in land surface phenology and greenness and to ensure the reliability of scientific results. Trend change detection methods can have a weak accuracy in



detecting the real underlying trend of a time series especially under high inter-annual variability. All methods were prone to detect trends opposite to the underlying reality. Phenology detection methods can result in large differences in the estimation of the start and end of growing season. Differences between satellite datasets further add to these methodical uncertainties. Therefore it is necessary to take observational uncertainties into account in the detection of dynamics in vegetation activity and in model-data integration.

The newly implemented phenology module in LPJmL with optimized model parameters outperforms the original model in reproducing several observed patterns and dynamics of vegetation activity such as the global patterns of gross primary production, biomass and tree cover; and the seasonal to decadal changes in vegetation greenness. Specifically, the model agreement with observations is higher than the agreement among satellite datasets regarding start and end of growing season dynamics in temperate, boreal and arctic ecosystems. Moreover, the improved model better reproduces observed trends in the seasonal amplitude of atmospheric CO<sub>2</sub> than the original model.

The comparison of LPJmL with observations suggests that the model can be used to diagnose controls of climate and different environmental conditions on vegetation greenness at multiple time scales. The importance of controls varies regionally and depends on time scale. However, water availability and temperature are co-dominant controls on vegetation greenness seasonality, inter-annual variability and trends globally. Water availability affects vegetation seasonality not only in water-limited biomes but also in boreal and arctic regions where it is regulated through the seasonal freezing and thawing of permafrost soils. An increase in water availability from increased seasonal melting of permafrost contributed to greening trends in the arctic tundra and boreal forests of Siberia. Additionally, changes in vegetation cover and composition through vegetation dynamics and land use change are important controls for the inter-annual variability and trends in land surface greenness globally.

The observed increase of the seasonal amplitude of atmospheric CO<sub>2</sub> in northern high latitudes can be mainly explained by a stronger increase in gross primary production rather than in ecosystem respiration. This results in an increasing carbon sink strength in boreal and arctic ecosystems. Although the effects of CO<sub>2</sub> fertilization on

photosynthesis and global agricultural areas contribute moderately to the trends in CO<sub>2</sub> amplitude, the latitudinal gradient, showing much stronger trends in northern latitudes than in low latitudes, cannot be explained without the climate-induced amplification of vegetation activity in boreal and arctic regions.

**General conclusions:** Overall the thesis demonstrates the value of satellite observations in understanding global vegetation dynamics, carbon cycle trends, and for the improvement of global ecosystem models. At the same time, observational uncertainties and multiple independent data streams need to be considered. A rigorous comparison of ecosystem observations and models provides a more comprehensive understanding of the functioning of global ecosystems and the Earth system. However, little is known about the global importance of ecosystem disturbances (beyond fire) and mortality on biogeographic patterns, ecosystem carbon turnover and thus the global carbon cycle. Compilations of ground observations such as plant traits or mortality events in harmonized global databases and new retrievals from upcoming multi-spectral and radar satellite missions of the land surface will need to be explored and persistently integrated with continuously developed ecosystem models to further uncover the role of ecosystems in the Earth system.

## Zusammenfassung

**Hintergrund:** Satelliten-Beobachtungen der Erdoberfläche zeigen großflächige Veränderungen der Vegetation seit den 1980er Jahren. Beispielsweise stieg weltweit die Vegetationsaktivität an („Ergrünen“ der Erde). Beobachtungen zeigen aber auch ein Absinken der Vegetationsaktivität („Erbraunen“) in einigen Regionen wie z. B. in Teilen der borealen Wälder oder in temperierten Grasländern. In temperierten und borealen Wäldern veränderte sich auch die Phänologie der Vegetation wie z. B. ein zeitigerer Beginn oder ein späteres Ende der Vegetationsperiode. Satelliten-basierte Indizes der Vegetationsaktivität sind Indikatoren für Veränderungen in der Photosynthese and damit im globalen Kohlenstoffkreislauf. So wurde beispielsweise vorgeschlagen, dass das Ergrünen und der Anstieg der Kohlenstoffaufnahme durch Photosynthese in nördlichen Ökosystemen den Anstieg der Amplitude des Jahresgangs von CO<sub>2</sub> in der Atmosphäre verursachen. Dieser Zusammenhang zwischen dem Ergrünen und der Saisonalität von atmosphärischem CO<sub>2</sub> wurde jedoch noch nicht umfassend untersucht im Vergleich mit möglichen anderen Ursachen wie einem Anstieg der landwirtschaftlichen Produktion.

Die Glaubwürdigkeit von Satelliten-Beobachtungen des Ergrünen und Erbrauens wird intensiv diskutiert, weil regionale Trends nicht in allen Satelliten-Datensätzen detektiert werden können und sich auch in Abhängigkeit von der verwendeten statistischen Analyseverfahren unterscheiden. Beispielsweise wurde das Erbraunen in borealen Wäldern in Nord-Amerika nur in einigen Datensätzen entdeckt und mehrere alternative Erklärungen für diese Trends existieren. Das generelle Ergrünen und der zeitigere Beginn der Vegetationsperiode in nördlichen Ökosystemen wurde meistens mit langfristig ansteigenden Temperaturen erklärt. Obwohl globale Ökosystemmodelle häufig auf Temperatur-Beziehungen basieren um Phänologie zu simulieren, können sie nicht gut saisonale bis langfristige (ca. 30 Jahre) Beobachtungen der Vegetationsentwicklung reproduzieren. Daher ist es notwendig, zusätzliche Faktoren in globalen Ökosystemmodellen zu berücksichtigen um beobachtete saisonale bis langfristige Veränderungen der Vegetationsaktivität zu reproduzieren und zu erklären.

**Ziele:** Das übergeordnete Ziel dieser Promotion ist es, ein besseres Verständnis von Veränderungen der Phänologie und der Vegetationsaktivität der Landoberfläche in den letzten Jahrzehnten zu schaffen wobei Satellitenbeobachtungen mit globalen Ökosystemmodellen integriert werden sollen. Dabei werden vier spezifische Forschungsziele adressiert (Abbildung 1.2 und Kapitel 1.5). Erstens soll bewertet werden, wie Unsicherheiten in den Beobachtungen zur Abschätzung von Veränderungen der Phänologie und der Vegetationsaktivität der Landoberfläche beitragen. Unsicherheiten ergeben sich aus Unterschieden zwischen verschiedenen Satellitendatensätzen und variablen Genauigkeiten von Zeitreihen-Analysemethoden, die genutzt werden um Trends, Veränderungen in Trends oder phänologische Ereignisse zu detektieren. Das zweite Ziel ist, die Simulation von Phänologie in einem globalen Ökosystemmodell zu verbessern (speziell in LPJmL, Lund-Potsdam-Jena managed Lands). Dabei soll bewertet werden wie sich diese Modellverbesserung auf Simulationen der Vegetationsaktivität und des Kohlenstoffkreislaufes auswirkt. Das dritte Ziel ist es, die regionale und globale Bedeutung von verschiedenen Einflussfaktoren auf saisonale, jährliche und langfristige Veränderungen der Phänologie und Vegetationsaktivität der Landoberfläche zu bewerten. Schließlich ist das vierte Ziel eine umfassende Bewertung des Beitrages von Vegetationsaktivität und -produktivität auf den Anstieg der saisonalen CO<sub>2</sub>-Amplitude in der Atmosphäre.

**Daten, Methoden und Modelle:** Eine Vielzahl von Datensätzen und Methoden wurde genutzt um die Forschungsfragen zu adressieren: Hinsichtlich der ersten Zielstellung wurden neue Satellitendatensätze der Vegetationsaktivität zusammen mit mehreren Zeitreihenanalysemethoden genutzt um die Beobachtungs-Unsicherheiten zu bewerten. Für die zweite Fragestellung wurde ein neues Phänologie-Modul in LPJmL (genannt LPJmL-GSI) implementiert und mehrere Beobachtungsdatensätze wurden in LPJmL integriert um diagnostische Modellsimulationen zu verbessern. Das neue Phänologiemodul berücksichtigt Effekte von Temperatur, Strahlung und Wasserverfügbarkeit auf die saisonale Blattentwicklung von Vegetation. Hinsichtlich der dritten Forschungsfrage wurde LPJmL in einem Modellexperiment genutzt um die regionale und globale Bedeutung von verschiedenen Einflussfaktoren auf die Vegetationsaktivität and Phänologie der Landoberfläche auf verschiedenen Zeitskalen zu bewerten. Die dabei berücksichtigten Faktoren sind Temperatur, Strahlung,

Wasserverfügbarkeit, Veränderungen der Landnutzung und -bedeckung, Waldbrände und CO<sub>2</sub>-Düngung der Photosynthese. Schließlich wurde LPJmL für die vierte Forschungsfrage zusammen mit einem atmosphärischen Transportmodell (TM3) genutzt um umfassend die Rolle verschiedener Faktoren, u. a. des Ergrünens von arktischen und borealen Ökosystemen, für Trends in der saisonalen Amplitude von CO<sub>2</sub> in der Atmosphäre zu erklären.

**Hauptergebnisse:** Diese Dissertation stellt eine umfassende Bewertung dar hinsichtlich, 1) der Einflussfaktoren für die Vegetationsaktivität auf saisonalen bis langjährigen Zeitskalen und 2) der Effekte von Veränderungen der Vegetation auf die saisonale CO<sub>2</sub>-Amplitude in der Atmosphäre.

Die Berücksichtigung und Bewertung von Unsicherheiten ist essentiell um Veränderungen der Phänologie und Vegetationsaktivität aus Satellitendaten robust zu detektieren und um damit die Glaubwürdigkeit wissenschaftlicher Ergebnisse sicherzustellen. Methoden der Trend- und Trendveränderungsanalyse können eine geringe Genauigkeit bezüglich der Detektion des realen Trends einer Zeitreihe aufweisen insbesondere dann, wenn die Zeitreihe eine hohe Jahr-zu-Jahr-Variabilität aufweist. In solchen Fällen können einige Methoden sogar Trends mit gegensätzlicher Veränderung als die realen Trends detektieren. Methoden zur Detektion von phänologischen Ereignissen können in enormen Unterschieden hinsichtlich des Beginns oder Endes der Vegetationsperiode resultieren. Unterschiede zwischen Satellitendatensätzen tragen zusätzlichen zur Unsicherheit bei. Daher ist es notwendig solche Unsicherheiten bei der Detektion von Veränderungen der Vegetationsaktivität und bei der Modell-Daten-Integration zu berücksichtigen

Das neue Phänologiemodul in LPJmL mit optimierten Modellparametern übertrifft das Originalmodell, da es einige beobachtete Muster und Dynamiken der Vegetationsaktivität besser repräsentiert, wie z. B. globale räumliche Muster der Bruttoprimärproduktion, der Biomasse und Baumbedeckung und zeitliche Dynamiken der Vegetationsaktivität, d. h. Saisonalität, Jahr-zu-Jahr-Variabilität und Trends. Die Übereinstimmung des Modells mit den Satellitenbeobachtungen ist insbesondere höher als die Übereinstimmung der Satellitenbeobachtungen untereinander hinsichtlich des Beginns und Endes der Vegetationsperiode in temperierten, borealen und arktischen Ökosystemen. Darüber hinaus reproduziert das Modell auch besser die

beobachteten Trends in der saisonalen CO<sub>2</sub>-Amplitude in der Atmosphäre als das Originalmodell.

Die Evaluierung von LPJmL anhand von Beobachtungsdaten schafft die Grundlage für die Modellanwendung um Einflussfaktoren für die Vegetationsaktivität auf verschiedenen Zeitskalen zu diagnostizieren. Die Bedeutung von Einflussfaktoren variiert dabei regional und hängt von der Zeitskala ab. Jedoch sind weltweit Wasserverfügbarkeit und Temperatur ko-dominante Faktoren für die Saisonalität, Jahr-zu-Jahr-Variabilität und für Trends der Vegetationsaktivität. Wasser limitiert die Vegetationsaktivität nicht nur in ariden und semi-ariden Regionen sondern auch in borealen und arktischen Regionen wo die Wasserverfügbarkeit durch das saisonale Gefrieren und Auftauen der obersten Schicht von Permafrostböden reguliert wird. Ein Anstieg der Wasserverfügbarkeit aufgrund des verstärkten saisonalen Auftauens des Permafrostes hat zum Ergrünen der arktischen Tundra und der borealen Wälder in Sibirien beigetragen. Zusätzlich sind Veränderungen der Vegetationsbedeckung durch natürliche Vegetationsdynamik und durch Landnutzungsänderungen bedeutende Faktoren für die Jahr-zu-Jahr-Variabilität und für Trends in der Vegetationsaktivität weltweit.

Der beobachtete Anstieg der saisonalen Amplitude von CO<sub>2</sub> in der Atmosphäre in nördlichen Breiten kann hauptsächlich durch einen stärkeren Anstieg der Brutto-Photosynthese im Vergleich zur Ökosystematmung und damit einem Anstieg der Kohlenstoffsенke in borealen und arktischen Ökosystemen erklärt werden. CO<sub>2</sub>-Düngung der Photosynthese und globale landwirtschaftliche Gebiete tragen mäßig zu CO<sub>2</sub>-Amplitudentrends bei. Jedoch kann der Gradient von stärkeren CO<sub>2</sub>-Amplitudentrends in nördlichen Breiten zu schwachen Trends in niedrigen Breiten nicht ohne den Effekt von Klimawandel auf die Verstärkung der Vegetationsaktivität in borealen und arktischen Regionen erklärt werden.

**Allgemeine Schlussfolgerung:** Die Dissertation demonstriert den Nutzen von Satellitenbeobachtungen für das Verständnis von globalen Dynamiken der Vegetation, von Trends im Kohlenstoffkreislauf und für die Verbesserung globaler Ökosystemmodelle. Dabei müssen Unsicherheiten der Beobachtungen und mehrere voneinander unabhängige Datensätze berücksichtigt werden. Ein konsequentes Gegenüberstellen von Ökosystembeobachtungen und -modellen erlaubt dabei

umfassendere Erklärungen über die Funktionsweise von globalen Ökosystemen und des Erdsystems. Jedoch ist das Wissen über die globale Bedeutung von Ökosystemstörungen (abgesehen von Waldbränden) und -mortalität für biogeographische Muster, für den Umsatz von Kohlenstoff in Ökosystemen und damit für den globalen Kohlenstoffkreislauf spärlich. Zusammenstellungen von *in situ* Beobachtungen wie z. B. von Pflanzenmerkmalen oder Mortalitätsereignissen in globalen harmonisierten Datenbanken und neue Abschätzungen der Vegetationsaktivität von zukünftigen multispektralen Satelliten bzw. der Biomasse von Radar-Satelliten müssen ständig in kontinuierlich fortentwickelte Ökosystemmodelle integriert werden, um die Rolle von Ökosystemen im Erdsystem weiter aufzudecken.

## Table of contents

<b>Contents</b> .....	5
<b>Summary</b> .....	6
<b>Zusammenfassung</b> .....	10
<b>Table of contents</b> .....	15
<b>Figures</b> .....	20
<b>Tables</b> .....	25
<b>List of manuscripts and publications</b> .....	26
Editorial note .....	26
Co-authored publications.....	27
<b>1 Introduction</b> .....	29
1.1 Vegetation greenness and climate: scope of the thesis .....	30
1.2 Greenness and phenology in the Earth system.....	35
1.2.1 Warming-greening-albedo feedback.....	36
1.2.2 Greening and the global carbon cycle .....	37
1.3 Remote sensing of greenness and phenology .....	42
1.3.1 Vegetation indices and biophysical properties .....	42
1.3.2 Observational uncertainties from sensors and datasets .....	44
1.3.3 Observational uncertainties from time series analysis methods .....	46
1.4 Model-data integration for greenness and phenology .....	48
1.4.1 Dimensions and groups of global ecosystem models .....	48
1.4.2 Evaluation of phenology and greenness in global ecosystem models .....	51
1.4.3 Concepts of model-data integration .....	53
1.4.4 Improved phenology modelling through model-data integration .....	56
1.5 Research questions .....	58
1.5.1 How do observational uncertainties contribute to the quantification of dynamics in land surface phenology and greenness?.....	58
1.5.2 How can DGVMs be improved to better represent phenology and greenness dynamics? .....	59
1.5.3 What are the controls on land surface phenology and greenness on multiple time scales?.....	60



1.5.4	What is the contribution of high latitude greening on the increasing amplitude of atmospheric CO <sub>2</sub> ?	61
<b>2</b>	<b>Trend change detection in NDVI time series: effects of inter-annual variability and methodology</b>	<b>63</b>
	Abstract	64
2.1	Introduction	65
2.2	Data and methods	68
2.2.1	GIMMS NDVI3g dataset	68
2.2.2	Breakpoint detection algorithm	69
2.2.3	Methods for trend estimation	70
2.2.4	Simulation of surrogate time series	72
2.2.5	Evaluation of method performances	77
2.2.6	Application to real time series of Alaska: ensemble of NDVI trends	79
2.3	Results	80
2.3.1	Observed and simulated properties of NDVI time series	80
2.3.2	Evaluation of estimated breakpoints	81
2.3.3	Evaluation of estimated trends	84
2.3.4	Effects on the overall performance of the methods	86
2.3.5	Multi-method ensemble of breakpoint and trend estimates in Alaska	90
2.4	Discussion	91
2.4.1	Effect of temporal resolution on method performance	91
2.4.2	Effect of trend changes on method performance	92
2.4.3	Effect of inter-annual variability on method performance	93
2.4.4	Plausibility of trend and breakpoint estimates in Alaska	94
2.4.5	Practical recommendations	97
2.5	Conclusions	99
	Acknowledgements	100
<b>3</b>	<b>Identifying environmental controls on vegetation greenness phenology through model-data integration</b>	<b>103</b>
	Abstract	104
3.1	Introduction	105
3.2	Model, data sets and model–data integration	108

3.2.1	Overview .....	108
3.2.2	FAPAR and phenology in the LPJmL DGVM.....	110
3.2.3	Data sets.....	114
3.2.4	Model–data integration .....	119
3.3	Results and discussion.....	124
3.3.1	Parameter optimization.....	124
3.3.2	Effects of an improved phenology module in LPJmL.....	130
3.3.3	Applicability and challenges of the LPJmL-GSI phenology module .....	137
3.3.4	Environmental controls on vegetation greenness phenology.....	139
3.4	Conclusions.....	143
	Acknowledgements.....	144
<b>4</b>	<b>Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness .....</b>	<b>147</b>
	Abstract.....	148
4.1	Introduction .....	149
4.2	Material and methods.....	152
4.2.1	FAPAR datasets.....	152
4.2.2	FAPAR and phenology in LPJmL.....	153
4.2.3	LPJmL model setup and factorial model experiments.....	154
4.2.4	Phenology methods and trend analysis.....	155
4.2.5	Data comparison, model evaluation and quantification of factorial effects	157
4.3	Results.....	159
4.3.1	Comparison of land surface phenology and greenness metrics from datasets and models.....	159
4.3.2	Controls on phenology and greenness dynamics.....	166
4.4	Discussion .....	170
4.4.1	Uncertainties in detection and explanation of land surface phenology and greenness .....	170
4.4.2	Importance of light, water availability, and LULCC for phenology dynamics.....	173
4.4.3	Regional controls for greening and browning trends.....	174

4.4.4	Controls for global greening trends and consequences for prognostic modelling.....	176
	Acknowledgments.....	177
<b>5</b>	<b>Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere .....</b>	<b>179</b>
	Abstract.....	180
	Main text.....	180
	Acknowledgements.....	189
<b>6</b>	<b>Conclusions and future research.....</b>	<b>191</b>
6.1	Accounting for observational uncertainties .....	192
6.2	From phenology modelling to multi-scale model-data integration.....	195
6.3	Co-dominant water control on phenology and greenness dynamics.....	199
6.4	Vegetation dynamics as regulator of carbon cycle trends.....	202
<b>7</b>	<b>Supplement of chapter 3.....</b>	<b>205</b>
7.1	LPJmL model details.....	206
7.1.1	Original phenology model (LPJmL-OP) .....	206
7.1.2	Water availability scaling factor.....	207
7.1.3	Albedo.....	208
7.2	FAPAR datasets.....	209
7.2.1	Comparison of the Geoland2 and GIMMS3g FAPAR datasets .....	209
7.2.2	Estimation of uncertainty for the GIMMS3g FAPAR dataset.....	211
7.3	Land cover.....	211
7.3.1	Creation of an observation-based map of plant functional types .....	211
7.3.2	Comparison of simulated and observed PFT distributions .....	214
7.4	Model parameter optimization .....	217
7.4.1	Parameter definitions and values .....	217
7.4.2	Genetic optimization algorithm .....	220
7.4.3	Parameter sensitivities and uncertainties .....	223
7.4.4	Supplementary results and discussion on optimization performance.....	224
7.4.5	Supplementary results and discussion on parameter variability.....	228
7.5	Global model evaluation.....	233
7.5.1	Supplementary results and discussion on carbon stocks and fluxes.....	233

7.5.2	Supplementary figures on evapotranspiration .....	236
7.5.3	Supplementary figures on evaluation of FAPAR .....	238
<b>8</b>	<b>Supporting information of chapter 4.....</b>	<b>241</b>
8.1	Description of phenology methods .....	242
8.2	Supporting figures .....	245
<b>9</b>	<b>Methods and extended data of chapter 5 .....</b>	<b>250</b>
9.1	Methods.....	251
9.1.1	CO <sub>2</sub> time series processing and site selection.....	251
9.1.2	Trend analysis, uncertainty estimation and latitudinal gradients .....	252
9.1.3	LPJmL: simulation of terrestrial carbon fluxes, agriculture and vegetation dynamics.....	252
9.1.4	TM3 atmospheric transport model, fossil fuel emissions, and net ocean CO <sub>2</sub> uptake.....	253
9.1.5	Factorial model experiments and quantification of contributions .....	254
9.2	Extended data tables and figures.....	256
	<b>References.....</b>	<b>265</b>
	<b>Acknowledgement / Danksagung / Poděkování.....</b>	<b>317</b>
	<b>Author contributions to the manuscripts .....</b>	<b>318</b>
	Erklärung zu den Eigenanteilen des Promovenden an den Publikationen und Zweitpublikationsrechten bei einer kumulativen Dissertation.....	318
	<b>Curriculum vitae.....</b>	<b>321</b>
	<b>Selbstständigkeitserklärung.....</b>	<b>323</b>

## Figures

Figure 1.1: A satellite’s view on the land surface of the Earth.....	29
Figure 1.2: Overview of research questions and the structure of this thesis. ....	33
Figure 1.3: The greening Earth - trends in annual peak NDVI (1982-2011). ....	34
Figure 1.4: Feedbacks between vegetation phenology and the Earth System.....	36
Figure 1.5: Simplified schematic of the global carbon cycle (Ciais et al., 2013, p.471)....	39
Figure 1.6: High-latitude greening and increasing amplitude of atmospheric CO <sub>2</sub> . ....	40
Figure 1.7: Reflectance of vegetation, soil and location of spectral bands from different satellite sensors. ....	43
Figure 1.8: Dimensions of global ecosystem models.....	49
Figure 1.9: Use of data in model development and model-data integration.....	53
Figure 1.10: Schematic diagram of the structure of a model-data integration framework. ....	54
Figure 2.1: Estimated time series components for a random-selected example grid cell in central Alaska (3*3 grid cells averaged around central pixel 146.424°W, 64.762°N)..	73
Figure 2.2: Spatial and statistical distributions of NDVI time series properties in Alaska and time series components of the simulated NDVI time series.....	74
Figure 2.3: Examples of simulated time series with different components of trend, IAV, seasonal and remainder referring to the simulated trend, inter-annual variability, seasonal and remainder time series components, respectively. ....	76
Figure 2.4: Frequencies of differences between real and estimated number of breakpoints for the different methods from all experiments.....	83
Figure 2.5: Distribution of the temporal absolute difference between real and estimated breakpoints.....	84
Figure 2.6: Comparison of real and estimated slopes from different methods, all time series segments and all experiments. ....	85
Figure 2.7: Distribution of the root mean square error (RMSE) between real and estimated trend component for the different methods from all experiments. ....	87
Figure 2.8: Ensemble of breakpoint and trend estimates from all methods in Alaska...	89
Figure 2.9: Comparison of detected breakpoints with temporal fire and climate patterns in Alaska. ....	95
Figure 2.10: Comparison of detected breakpoints of the year 2004 from four different methods with 2004 burnt areas and <i>in situ</i> photos (taken in 2008). ....	95
Figure 3.1: Structure of the model–data integration approach for LPJmL (LPJmL-MDI). ....	108

Figure 3.2: Examples of the cold temperature, heat stress, light and water limiting functions for phenology in LPJmL-GSI.....	112
Figure 3.3: Map of the dominant PFT in each grid cell as derived from SYNMAP, Köppen–Geiger climate zones and MODIS VCF.....	119
Figure 3.4: Distribution of the correlation coefficient between monthly LPJmL and GIMMS3g FAPAR (1982–2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes.....	125
Figure 3.5: Uncertainty and sensitivity of LPJmL-GSI parameters derived from all individuals of genetic optimizations at PFT level. ....	127
Figure 3.6: Best LPJmL model runs for (a) monthly FAPAR dynamics (1982–2011, n = 360 months) and (b) time series of mean annual FAPAR (1982–2011, n = 30 years).....	132
Figure 3.7: Comparison of trends in mean annual FAPAR from LPJmL and from satellite data. ....	133
Figure 3.8: Latitudinal gradients of (a) gross primary production (GPP), (b) evapotranspiration, (c) biomass and (d) tree cover from data-oriented estimates and from LPJmL model simulations. ....	135
Figure 3.9: Phenological controls on seasonal FAPAR dynamics. ....	140
Figure 4.1: Flow chart of the described datasets and methods. ....	152
Figure 4.2: Latitudinal gradients of mean annual (a) SOS, (b) EOS and (c) peak FAPAR from datasets and LPJmL.....	160
Figure 4.3: Regional averaged time series and linear trends of (a-g) peak FAPAR and (h-i) SOS.....	162
Figure 4.4: Kling-Gupta efficiency for SOS, EOS and peak FAPAR for datasets and LPJmL phenology models.....	163
Figure 4.5: Kling-Gupta efficiency (KGE) for satellite data comparison and LPJmL-GSI model evaluation for peak FAPAR (a-h) and SOS (i-k) globally and for different regions.....	164
Figure 4.6: Total effect of different factors on (a) peak FAPAR, (b) SOS and (c) EOS..	165
Figure 4.7: Latitudinal gradients of the total effect of different factors on peak FAPAR. ....	166
Figure 4.8: Regional distributions and averaged total effects with components on (a-g) peak FAPAR and (h-i) SOS. ....	167
Figure 4.9: Global temporal dynamic and controlling factors for land surface phenology and greenness. ....	169
Figure 5.1: Amplification of plant activity in the northern biosphere. ....	183
Figure 5.2: Latitudinal gradients of trends in the seasonal CO <sub>2</sub> amplitude and its drivers. ....	184

Figure 5.3: Trends in the seasonal amplitude of terrestrial net biome productivity (NBP), gross primary productivity (GPP) and ecosystem respiration (Reco) per latitudinal bands of 3° width. ....	186
Figure 7.1: Effects on FAPAR in LPJmL for an example grid cell in Siberia.....	207
Figure 7.2: Standard deviation of mean annual FAPAR from the GIMMS3g and GL2 FAPAR datasets in 1982-2011. ....	209
Figure 7.3: Comparison of mean annual FAPAR from different datasets averaged for the extent of boreal needle-leaved evergreen forests. ....	210
Figure 7.4: Monthly quantile regressions between GL2 VGT FAPAR and the GL2 VGT FAPAR fitted to the quantile 0.95. ....	211
Figure 7.5: Reclassification of the Koeppen-Geiger climate classification in bioclimatic zones.....	212
Figure 7.6: Comparison of total forest coverage from SYNMAP and MODIS tree coverage for a region in eastern Siberia. ....	214
Figure 7.7: Observation-based maps of the foliar projective cover of plant functional types. Agricultural areas are included in the TrH and TeH PFTs.....	215
Figure 7.8: Comparison between simulated and observed PFT distributions for the year 2000.....	217
Figure 7.9: Information sources for prior and posterior parameter sets and overview of model optimization experiments. ....	218
Figure 7.10: Distribution of the cost for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes.....	225
Figure 7.11: Distribution of the percent bias between LPJmL and MTE mean annual GPP (1982-2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes. See Figure 7.10 for a further explanation of this figure.....	226
Figure 7.12: Distribution of the percent bias between LPJmL and MODIS monthly growing season albedo (2000-2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes. See Figure 7.10 for a further explanation of this figure. ....	228
Figure 7.13: Prior and optimized values for the parameter $\alpha$ (fraction of radiation absorbed at leaf level relative to canopy level) grouped by plant functional types and biomes. ....	229
Figure 7.14: Prior and optimized values for the parameter $\beta_{\text{leaf}}$ (leaf albedo) grouped by plant functional types and biomes.....	230
Figure 7.15: Prior and optimized values for the parameter $k$ (light extinction coefficient) grouped by plant functional types and biomes. ....	231
Figure 7.16: Correlations between posterior parameters for the four limiting functions for phenology in LPJmL-GSI grouped per PFT. ....	232

Figure 7.17: Comparison of patterns of mean annual total gross primary production from LPJmL and the data-oriented MTE estimate for the period 1982-2011.....	234
Figure 7.18: Comparison of the mean seasonal GPP cycle (averaged over 1982-2011) from MTE and LPJmL spatially averaged for regions with the same dominant PFT..	235
Figure 7.19: Comparison of biomass from data-oriented estimates (Thurner and Saatchi datasets) and from LPJmL (averaged 2009-2011). See Figure 7.17 for further explanations. ....	235
Figure 7.20: Latitudinal gradients of evapotranspiration with its components. ....	236
Figure 7.21: Mean seasonal cycle (1982-2011) of ET, evaporation, interception and transpiration spatially averaged for PFTs.....	237
Figure 7.22: Comparison of mean annual FAPAR from LPJmL and remote sensing datasets. ....	238
Figure 7.23: Comparison of the mean seasonal FAPAR cycle from GIMMS3g, GL2 VGT and LPJmL spatially averaged for regions with the same dominant PFT. ....	238
Figure 7.24: Correlation coefficients between monthly FAPAR time series from GIMMS3g, GL2 VGT datasets and LPJmL model simulations.....	239
Figure 7.25: Correlation coefficients between annual FAPAR time series from GIMMS3g, GL2 VGT datasets and LPJmL model simulations.....	239
Figure 7.26: Extrapolation capabilities of LPJmL-GSI in terms of monthly FAPAR dynamics.....	240
Figure 8.1: Used masks for (a) biome-aggregated, (b) region-aggregated and (c) global-aggregated results in tables and figures of the main text.....	245
Figure 8.2: Latitudinal gradients of all PGMs. See Figure 4.2 of the main text for an explanation.....	246
Figure 8.3: Regional-averaged time series and trends in SOS, EOS and LOS. See Figure 3 of the main text for an explanation. ....	247
Figure 8.4: Latitudinal gradients of median annual SOS as estimated from 10 phenology methods for (a) GIMMS3g, (b) MODIS and (c) GL2-VGT2.....	247
Figure 8.5: RGB composite maps of all effects on all PGMs.....	248
Figure 8.6: KGE between SOS from LPJmL-GSI and the data-method ensemble.....	249
Figure 8.7: Globally spatial averaged time series and trends of (a) peak FAPAR, (b) water availability anomalies from LPJmL, (c) anomalies in total tree foliar projective cover from LPJmL and (d) air temperature and precipitation anomalies. All anomalies are relative to 1982.....	249
Figure 9.1: Uncertainty in observed CO <sub>2</sub> amplitude trends at (a) BRW and (b) MLO depending on the length of the time series.....	261
Figure 9.2: Changes in CO <sub>2</sub> amplitude at 500 mb. ....	261
Figure 9.3: Effect of different meteorological forcing and ocean CO <sub>2</sub> uptake datasets on CO <sub>2</sub> amplitude trends. ....	262



Figure 9.4: Evaluation of simulated (a) mean values and (b) trend slopes of CO<sub>2</sub> amplitude against observations. .... 263

Figure 9.5: Latitudinal gradients of trends of (a) the area coverage of plant functional types as simulated with LPJmL, (b) annual maximum vegetation greenness (FAPAR, fraction of absorbed photosynthetic active radiation), (c) annual amplitude of gross primary production (GPP), and (d) of net biome productivity (NBP). .... 263

Figure 9.6: Effects of (a) terrestrial gross fluxes, (b) regions, and (c) LPJmL model components on the latitudinal gradient of CO<sub>2</sub> amplitude trends..... 264

## Tables

Table 2.1: Normalized confusion matrices of estimated and real trend classes for each method.....	86
Table 2.2: Analysis of variance table for the RMSE between real trend and estimated trend. IAV and STV denote inter-annual and short-term variability, respectively. NBP is the number of breakpoints.....	88
Table 3.1: Overview of optimization experiments with information sources for prior and posterior parameter sets. Parameter values and prior parameter ranges for each parameter set are listed in Supplement 7.4.1.....	123
Table 3.2: Overview of global model runs that were used in this study for model evaluation.....	124
Table 4.1: Overview of factorial model experiments and corresponding effects on FAPAR in LPJmL.....	155
Table 4.2: Biome-averaged mean annual start of season (SOS) and end of season (EOS) (in days of year, DOY) from satellite-based datasets (GIMMS3g, MODIS, GL2-VGT2) and LPJmL phenology models.....	160
Table 4.3: Biome-averaged trends in SOS (days/year).....	161
Table 7.1: Description of LPJmL model parameters that were addressed in this study.....	219
Table 7.2: Prior parameter values of LPJmL-OP (OP.prior).....	220
Table 7.3: Posterior parameter values for LPJmL-OP based on grid cell-level optimization experiments (OP.gc).....	221
Table 7.4: Prior parameter values for LPJmL-GSI (GSI.prior).....	222
Table 7.5: Final parameters for LPJmL-GSI.....	223
Table 7.6: Global total carbon fluxes and stocks from data-oriented estimates and from LPJmL simulations.....	234
Table 8.1: Explanation of abbreviations.....	242
Table 9.1: Summary information about the used CO <sub>2</sub> measurement stations.....	256
Table 9.2: Overview of factorial model experiments with LPJmL+TM3.....	257
Table 9.3: Changes in the contribution to the seasonal amplitude of CO <sub>2</sub> separated (a) by regions and (b) by terrestrial gross fluxes.....	259

## List of manuscripts and publications

The following manuscripts or publications are part of this cumulative doctoral thesis:

Chapter 2: Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M., Neigh, C. and Reichstein, M.: Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology, *Remote Sensing*, 5(5), 2113–2144, doi:10.3390/rs5052113, 2013.

Chapter 3: Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

Chapter 4: Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U. and Carvalhais, N.: Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness, submitted to *Global Change Biology*, 2014-12-16.

Chapter 5: Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R. F., Heimann, M., Thonicke, K., Zaehle, S. and Reichstein, M.: Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere, submitted to *Nature*, 2015-01-30.

### Editorial note

All publications and manuscripts were edited to obtain a common style and format in this doctoral thesis. Editorial modifications were made regarding: 1) fonts, and layout of tables; 2) numbering of chapters, figures, and tables. 3) The citation style system of the journal *Biogeosciences* was used for all manuscripts and all references were combined in a single reference list. 4) Supplements or extended data of the respective articles were included as chapters 7, 8 and 9. Nevertheless, text, figures, and contents of tables fully agree with the original publications.

## Co-authored publications

The following publications were authored or co-authored during the doctoral studies but are not included in this doctoral thesis:

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T. and Reichstein, M.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, *Nature*, 514(7521), 213–217, doi:10.1038/nature13731, 2014.

Forkel, M., Thonicke, K., Beer, C., Cramer, W., Bartalev, S. and Schmulius, C.: Extreme fire events are related to previous-year surface moisture conditions in permafrost-underlain larch forests of Siberia, *Environ. Res. Lett.*, 7(4), 044021, doi:10.1088/1748-9326/7/4/044021, 2012.

Semenova, O., Lebedeva, L., Volkova, N., Korenev, I., Forkel, M., Eberle, J. and Urban, M.: Detecting immediate wildfire impact on runoff in a poorly-gauged mountainous permafrost basin, *Hydrological Sciences Journal*, doi:10.1080/02626667.2014.959960, 2014.

Urban, M., Forkel, M., Eberle, J., Hüttich, C., Schmulius, C. and Herold, M.: Pan-Arctic Climate and Land Cover Trends Derived from Multi-Variate and Multi-Scale Analyses (1981–2012), *Remote Sensing*, 6(3), 2296–2316, doi:10.3390/rs6032296, 2014.

Urban, M., Forkel, M., Schmulius, C., Hese, S., Hüttich, C. and Herold, M.: Identification of land surface temperature and albedo trends in AVHRR Pathfinder data from 1982 to 2005 for northern Siberia, *International Journal of Remote Sensing*, 34(12), 4491–4507, doi:10.1080/01431161.2013.779760, 2013.



# 1 Introduction

---

*“Wenn auch der Charakter der verschiedenen Erdräume von allen äußeren Erscheinungen abhängt (...) so ist doch nicht zu leugnen, dass das Hauptbestimmende dieses Eindrucks die Pflanzendecke ist.”*

(Humboldt 1845/2004, p.180)

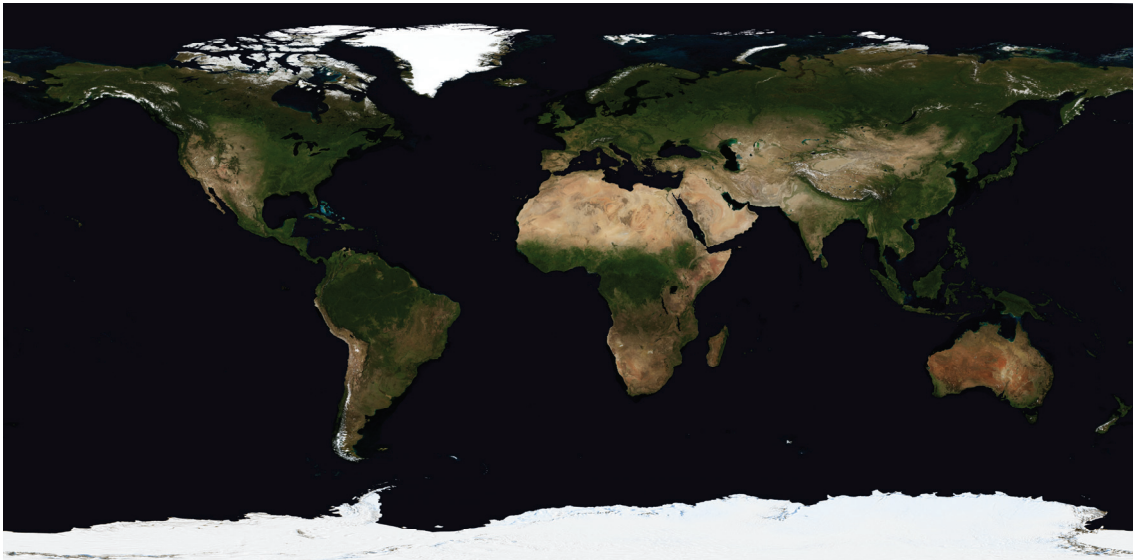


Figure 1.1: A satellite's view on the land surface of the Earth.

This image is a cloud-free true-colour composite of MODIS satellite observations from July 2004 from the NASA Blue Marble Next Generation dataset (Stöckli et al., 2005).

# **1 Introduction**

## **1.1 Vegetation greenness and climate: scope of the thesis**

Alexander von Humboldt, the pioneer of modern physical geography, had already emphasized the dominance of vegetation on the appearance of different regions (Humboldt 1845/2004, p.180). Humboldt describes that the presence of certain plant groups changes from the equator to the poles or from low to high latitudes. This change in plant groups is indicative for differences in climate conditions (Humboldt 1845/2004, p.182). The close relationships between the distribution of plant species and climate have been later used by Köppen (1884) and Köppen and Geiger (1954) to define and map climate zones that correspond to dominant plant species. The definition of these climate zones considered also seasonal variations of temperature and precipitation. The seasonality of temperature and precipitation has later been described and especially visualized in climographs by Walter and Lieth (1960). Thereby climographs were mostly developed based on temperature and water availability conditions that are descriptive of the seasonality of plant growth (Walter, 1970). The seasonality of plant growth is one of the most noticeable phenomena in nature that shows the strong interaction between climate and vegetation: new leaves of deciduous plants emerge in spring and fall again in autumn. The study of the timing of such seasonal biological changes is called phenology (Lieth, 1976). Phenology of vegetation is usually controlled by temperature, day length and dormancy (Körner and Basler, 2010; Lieth, 1974). Therefore phenology of vegetation has a strong year-to-year variability following variations in weather conditions (Walther et al., 2002). For example, annual weather conditions and harvest yields have been documented over centuries in monasteries, which reported a close relationship between the variability of air temperature and vine harvest dates in Europe (Chuine et al., 2004; Maurer et al., 2009). Modern phenological observations demonstrate an advanced leafing, flowering or fruiting in Europe in the last four decades which is caused by warming (Menzel et al., 2006). Thus, changes in the phenology and in distribution of the terrestrial vegetation are direct indicators for past and recent changes in climate conditions.

Satellite observations of the Earth are nowadays indispensable tools to investigate interactions between climate and the terrestrial vegetation (Figure 1.1). The emergence of aerial photography during the first half of the 20<sup>th</sup> century and of Earth observation satellites during the 1970s opened possibilities to complement local ground observations of phenological changes with spatial covered observations (Morain, 1974). Nevertheless, the use of satellite images to study phenology required spectral observations that are comparable under different vegetation types and temporal and spatial conditions. Only the development of spectral vegetation indices like the NDVI (Normalized Difference Vegetation Index, chapter 1.3.1) allowed to observe temporal changes in vegetation conditions (Tucker, 1979). Vegetation indices are measures of the “greenness” of the vegetation (Huete, 2012) and are based on the characteristic spectral properties of vegetation. Vegetation reflects relatively more radiation in the green region (ca. 500-600 nm) of the visible part of the electromagnetic spectrum by chlorophyll than in the blue (ca. 400-500 nm) and the red regions (ca. 600-700 nm). Radiation in blue and red regions is absorbed as energy source for photosynthesis (Emerson and Lewis, 1943; Engelmann, 1894). Additionally, almost all incoming solar radiation is reflected in the near-infrared region. Thus, the difference between the high reflection (low absorption) in the near-infrared and the low reflection (high absorption) in the red region is a proxy for the photosynthetic light absorption. This difference usually increases with an increasing coverage of green vegetation. Therefore the value of a vegetation index is a measure of the “greenness” although spectral vegetation indices are based on the absorption of radiation in the red region.

Multi-temporal satellite observations of vegetation indices made it possible to study the spatial variability and temporal development of greenness (Justice et al., 1985) and phenology (Reed et al., 1994). The use of satellite-derived time series of vegetation indices to study variations in vegetation phenology is usually referred to as “land surface phenology” because satellite observations integrate signals of the land surface across different ecosystems and species (de Beurs and Henebry, 2004a, 2010b). Satellite observations during the last three decades allow the derivation of vegetation index time series that can be used to detect changes in land surface phenology and greenness. Time series of vegetation indices show significant changes in land surface phenology and greenness since the 1980s. Already in the period from 1981 to 1991 an increased



plant growth measured by an increase in annual growing season NDVI has been observed in the Northern Hemisphere (Myneni et al., 1997a). Such positive trends in time series of vegetation indices are called “greening” and occur in many terrestrial regions across the Earth (Figure 1.3). Greening trends in high latitudes are accompanied by trends towards an earlier start and a lengthening of the growing season (Julien and Sobrino, 2009; Tateishi and Ebata, 2004; Tucker et al., 2001). These greening and growing season trends in high-latitude ecosystems can be explained by increasing temperatures that enhance plant growth (Keenan et al., 2014; Lucht et al., 2002; Xu et al., 2013). Greening trends occur also in the Sahel (Herrmann et al., 2005; Olsson et al., 2005) and are explained by a recovery of vegetation after severe droughts in the 1960s and 1970s (Hickler et al., 2005). Nevertheless, the greening of the Sahel is highly debated because it cannot be detected in all satellite datasets (Fensholt and Proud, 2012; Horion et al., 2014) and it is generally surprising given the fact that the Sahel has experienced strong land degradation in the last decades (Dardel et al., 2014b; Fensholt et al., 2013). Besides greening, also negative trends in vegetation greenness (named “browning”) have been observed, especially in boreal forests of North America (Bi et al., 2013; Goetz et al., 2005) (Figure 1.3). Browning trends in boreal forests occur more often in evergreen than in deciduous forests and more often in North America than in Siberia (Beck and Goetz, 2011). In newer NDVI datasets with extended time series after 2006 a change from greening to browning trends has also been observed in some boreal forests of Eurasia (Buermann et al., 2014; de Jong et al., 2011b) (Figure 1.3). Browning trends in boreal forests have been explained by several and often regionally different factors like increased fire activity (Goetz et al., 2005), increasing drought stress (Buermann et al., 2014; Bunn et al., 2007), cooling spring temperatures (Wang et al., 2011) or decreased soil moisture because of reduced snow packs (Barichivich et al., 2014). Browning trends have also been observed in some temperate and subtropical grasslands in Asia and South America (de Jong et al., 2011b) which are explained by an increased drought stress and increased pasture farming (Hilker et al., 2014; van Leeuwen et al., 2013). Greening, browning, and phenology trends in boreal and arctic regions in the last three decades are seen as first signs of climate change induced shifts in the global distribution of vegetation (Beck et al., 2011). It is hypothesized that the boreal forest biome shifts northwards at the expense of the tundra biome given climate

warming scenarios for the next century (Lucht et al., 2006). The existence of multiple competing explanations for greening and browning trends in biomes such as boreal forests or the Sahel as well as the interpretation of these trends as first signs of biome shifts require a better understanding of the climatic, environmental and anthropogenic controls on dynamics of land surface phenology and greenness.

The overall aim of this thesis is to *better understand recent dynamics of land surface phenology and greenness in the Earth system* by combining satellite observations with global vegetation models. The introduction continues with a state-of-the-art assessment about the role of vegetation greenness and phenology in the Earth system (chapter 1.2) and with two methodological chapters focussed on remote sensing (chapter 1.3), modelling, and model-data integration (chapter 1.4, Figure 1.2). The introduction ends with a description of the four major research questions of this thesis (chapter 1.5). The research questions are addressed in several research chapters (chapters 2-5 with corresponding supplements in chapters 7-9). The main conclusions of the thesis as a whole are summarized with respect to the research questions in chapter 6.

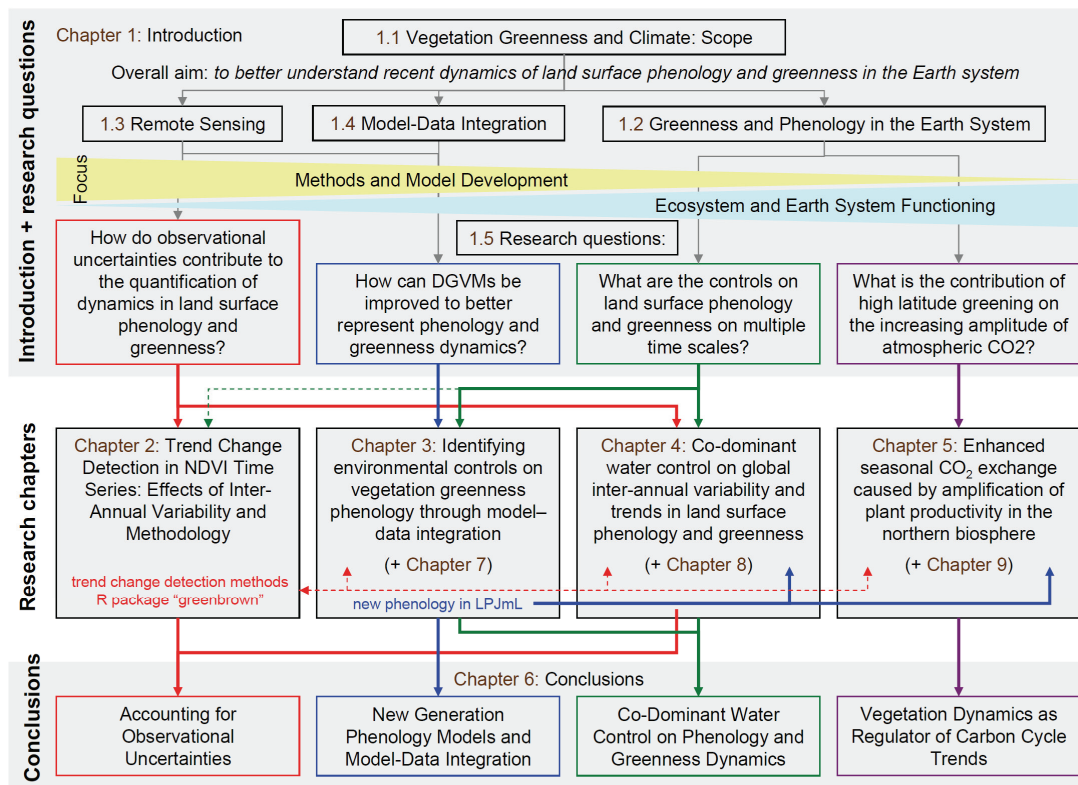


Figure 1.2: Overview of research questions and the structure of this thesis.

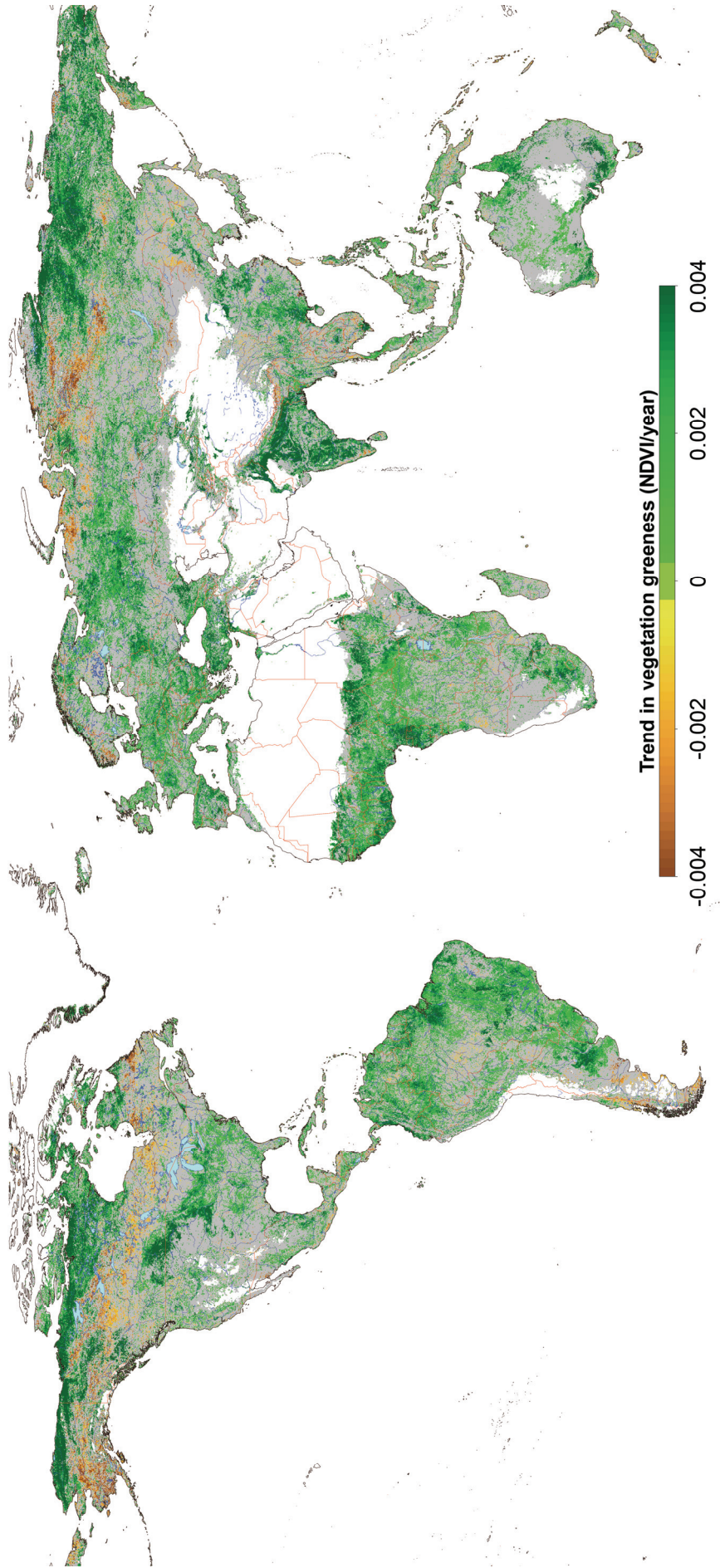


Figure 1.3: The greening Earth - trends in annual peak NDVI (1982-2011). Trends were calculated on the GIMMS3g NDVI dataset (Pinzon and Tucker, 2014). Method AAT-peak as described in chapter 2 was used to calculate trends by considering at maximum one breakpoint. The trend of the longest time series segment is shown in this map. Non-significant trends ( $p > 0.05$ , Mann-Kendall test) are coloured in grey.

## 1.2 Greenness and phenology in the Earth system

The greenness of vegetation is regulated through the temporal development and senescence of foliage, the physiological activity of canopies and the plant composition of ecosystems (Figure 1.4) (Richardson et al., 2013). Thus, the seasonal dynamic of vegetation greenness is to a large extent regulated through vegetation phenology. Vegetation greenness and phenology affect the Earth system by several mechanisms (Figure 1.4) (Bonan, 2008a, 2008b; Peñuelas et al., 2009; Richardson et al., 2013). These mechanisms are regulated by a structural and a physiological component of vegetation phenology (Richardson et al., 2013). The obvious structural component of phenology through leaf development and senescence regulates litterfall, microclimate, surface roughness and albedo. Albedo is the fraction of the incident solar radiation that is reflected by the surface (Bonan, 2008a, p.199). Albedo regulates the energy balance of the land surface and thus affects also the energy balance of the atmosphere. Green vegetated land surfaces are associated with low albedo and therefore increased absorbed radiation. An increased absorbed radiation alters the surface energy balance by increasing latent heat fluxes, meaning higher evapotranspiration and therefore increased water cycling. In a modelling scenario, contrasting a “desert” with a “green” Earth, Kleidon et al. (2000) estimated that the maximum potential effect of the terrestrial vegetation is to triple land evapotranspiration. Thus the structural development of vegetation canopies regulates the physiological activity of vegetation and controls photosynthesis, evapotranspiration and energy fluxes (Figure 1.4). These phenology-regulated mechanisms influence the composition of the atmosphere through the uptake of carbon dioxide (CO<sub>2</sub>) by photosynthesis, through the release of water (H<sub>2</sub>O) by evapotranspiration or the emission of volatile organic compounds (VOCs) (Peñuelas et al., 2009). The biogeophysical effect of land surface albedo and the biogeochemical effect of photosynthetic carbon uptake are thought to be the dominant regulators of the green terrestrial vegetation that affect the Earth system.

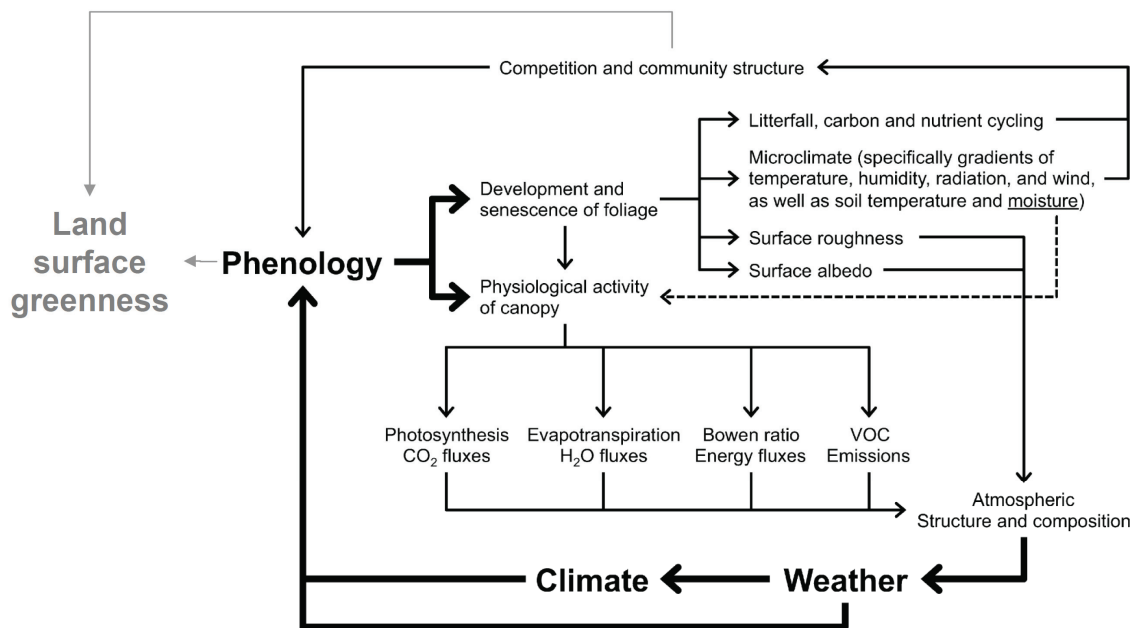


Figure 1.4: Feedbacks between vegetation phenology and the Earth System. The scheme was adapted from Richardson et al. (2013). Grey-coloured text and arrows were added to the original figure.

### 1.2.1 Warming-greening-albedo feedback

Greening trends and associated changes in land surface albedo could potentially contribute to further climate warming especially in arctic regions (Chapin et al., 2005). Land surface albedo is strongly affected by vegetation type. For example, evergreen needle-leaved forests have the lowest albedo values (0.05-0.15), followed by deciduous broadleaved forests (0.15-0.2), and grass or croplands (0.16-0.26) (Bonan, 2008a, p.200; Cescatti et al., 2012; Hollinger et al., 2010). The phenology of these vegetation types affects strongly the seasonality of land surface albedo. Whereas evergreen forests have not or only a small seasonal variation of albedo because of the stable canopy structure, the seasonal development of foliage in grasslands and deciduous forests cause larger variations in albedo (e.g. seasonal amplitude of albedo ~0.05 in deciduous forests, Cescatti et al., 2012). Additionally, the seasonal development of canopies affects the relative contribution of background albedo from litter, soil or snow cover to the total land surface albedo. Snow has high albedo values (> 0.45 for old snow, > 0.8 for fresh snow) (Bonan, 2008a, p.200). Thus, the albedo seasonality of seasonal snow-covered ecosystems is mostly affected by snow melt and snow fall.

Greening trends in high-latitude ecosystems, especially in the arctic tundra, are directly connected with changes in land surface albedo (Chapin et al., 2005). The arctic greening

is related to an increase of shrub cover (Forbes et al., 2010; Myers-Smith et al., 2011; Sturm et al., 2001) which is mainly due to warming and thus more favourable growth conditions (Blok et al., 2011a). An increasing shrub cover is in turn related to an increasing NDVI and decreasing land surface albedo (Loranty et al., 2011). A further increase of vegetation cover in arctic regions will decrease land surface albedo and consequently result in an increased absorption of solar radiation. This increased absorption will result in surface warming and an increased heat flux to the atmosphere which could amplify climate warming (Chapin et al., 2005; Jeong et al., 2014; Myers-Smith et al., 2011). Indeed, positive relations between land surface temperature and greenness, and negative relations between greenness and albedo have been observed in the taiga-tundra transition zone in the last three decades (Urban et al., 2013). On the other hand it has been also shown that increasing shrub cover does not generally decrease land surface albedo, e.g. in cases when shrubs replace wetlands with initial low albedo (Blok et al., 2011b). Nevertheless, an increasing coverage of shrubs or trees generally decreases land surface albedo especially during the snow season when the trees or shrubs are masking the underlying snow (Loranty et al., 2014). Consequently, the increasing tree and shrub cover in high latitude regions that can be observed as greening, decreases land surface albedo especially in winter and therefore contributes to increasing winter temperatures (Chapin et al., 2005; Loranty et al., 2014). Thus the warming is causing arctic greening and decreasing land surface albedo, meaning increased absorption and surface warming. Therefore high-latitude greening and associated albedo changes are part of a positive feedback in the Earth system.

## **1.2.2 Greening and the global carbon cycle**

### *1.2.2.1 Overview of the global carbon cycle*

Contrary to the greening-albedo feedback, the relation between greening and photosynthesis could dampen climate warming. Plants assimilate CO<sub>2</sub> from the atmosphere during photosynthesis. Global terrestrial gross photosynthesis (or gross primary production, GPP) is the largest flux of CO<sub>2</sub> in the carbon cycle (Figure 1.5) with 123 PgC yr<sup>-1</sup> globally with an uncertainty of 8 PgC yr<sup>-1</sup> (standard deviation across different estimate approaches) (Beer et al., 2010). Part of this assimilated carbon is respired as energy source for plant growth and to maintain living cells. The sum of this

growth and maintenance respiration, autotrophic respiration ( $R_a$ ), is about 50% of GPP (Bonan, 2008a, p.275). The difference between GPP and  $R_a$  is referred to as net photosynthesis (or net primary production, NPP) and is about 56 PgC yr<sup>-1</sup> with an uncertainty of 14.3 PgC yr<sup>-1</sup> (standard deviation across different estimates) (Ito, 2011). This carbon is used by plants to produce biomass. Around 450-650 PgC are stored as biomass at the global land surface (Ciais et al., 2013) (Figure 1.5). Carbon is transferred from vegetation to soils through mortality and litterfall which is around 32% to 43% of NPP (Bonan, 2008a, p.310). Soils are the largest terrestrial carbon stock with 1500-2400 PgC with additionally approx. 1700 PgC stored in permafrost (Ciais et al., 2013) (Figure 1.5). The large uncertainties in soil carbon stock estimates depend on the limited representativeness of soil profile measurements (Carvalhais et al., 2014). Carbon in litter and soils is released back into the atmosphere through respiration by microorganisms (heterotrophic respiration,  $R_h$ ). Carbon in ecosystems is also released to the atmosphere through fires (1.71-2.46 PgC yr<sup>-1</sup>, Thonicke et al. (2010)) or exported by rivers to oceans (~ 1.7 PgC yr<sup>-1</sup>, Ciais et al. (2013)). As GPP and total ecosystem respiration ( $Reco = R_a + R_h$ ) are the largest carbon fluxes in ecosystems, the net carbon uptake of an ecosystem (net ecosystem productivity, NEP, or net ecosystem exchange, NEE), is in most ecosystems dominated by these two components (Chapin et al., 2006; Schulze, 2006):

$$NEP = GPP - Reco = (GPP - R_a) - R_h = NPP - R_h \quad (1.1)$$

$$NEE = Reco - GPP = -NEP \quad (1.2)$$

The entire carbon balance of an ecosystem includes also non-respiratory losses such as fire C emissions, emissions of volatile organic compounds, and lateral transport and is defines as net ecosystem carbon balance (NECB) (Chapin et al., 2006) or for large scales as net biome productivity (NBP) (Schulze, 2006). The global terrestrial NBP is around 2.6 PgC yr<sup>-1</sup> (Ciais et al., 2013) (Figure 1.5), i.e. the terrestrial land surface takes up CO<sub>2</sub> and thus is a “sink” for atmospheric CO<sub>2</sub>. The amount of carbon that is stored in the atmosphere increases by ca. 4 PgC yr<sup>-1</sup> because of increased emissions from burning of fossil fuels and cement production (7.8 PgC yr<sup>-1</sup>) and changes in land use (1.1 PgC yr<sup>-1</sup>) (Ciais et al., 2013) (Figure 1.5). Similar in magnitude to terrestrial NBP, oceans take up atmospheric carbon (2.3 PgC yr<sup>-1</sup>) and store it mainly in the deep ocean (37100 PgC) (Ciais et al., 2013) (Figure 1.5). Nowadays, terrestrial ecosystems take up

more carbon ( $+1.2 \text{ PgC yr}^{-1}$ , Ciais et al. (2013)) than before the industrial revolution (~1750) which is largely driven by stronger trends in gross primary production than in ecosystem respiration. In summary, the terrestrial land surface is a net sink of carbon, dampening the growth of atmospheric  $\text{CO}_2$  concentrations which is a negative feedback in the Earth system.

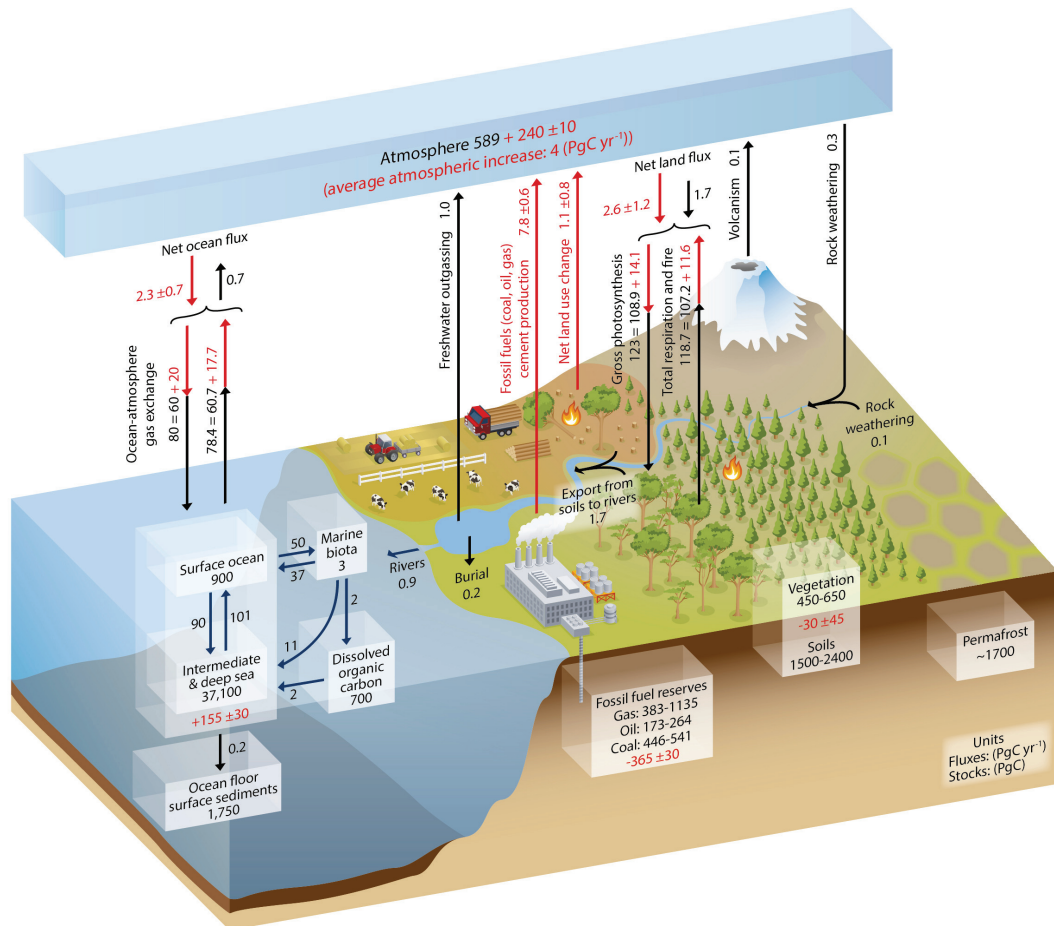


Figure 1.5: Simplified schematic of the global carbon cycle (Ciais et al., 2013, p.471). Numbers represent reservoir mass, also called ‘carbon stocks’ in  $\text{PgC}$  and annual carbon exchange fluxes (in  $\text{PgC yr}^{-1}$ ). Black numbers and arrows indicate reservoir mass exchange fluxes estimated for the time period prior to the industrial era, about 1750 (see Ciais et al. (2013, p.471) for a further description).

### 1.2.2.2 Greening and carbon cycle trends

Greening trends are concordant to positive trends in photosynthetic carbon uptake. Greening trends in high latitudes in 1982-2000 caused increased absorption of radiation for photosynthesis and thus resulted in positive NPP trends (Lucht et al., 2002; Nemani et al., 2003; Sitch et al., 2007). Lucht et al. (2002) estimated an annual NPP increase by



34.6 gC m<sup>-2</sup> yr<sup>-1</sup> in the boreal zone in 1982-1998 by using the LPJ model. A relationship between greening and increased NPP in the taiga-tundra transition zone is also confirmed by ground observations that demonstrate significant correlations between annual mean growing season NDVI and annual tree ring widths (Berner et al., 2011, 2013). Nemani et al. (2003) estimated a global increase of NPP of 3.42 PgC in the period 1982-1999 which was mostly dominated from tropical regions. Nevertheless, for the global terrestrial carbon budget Reco needs to be considered to quantify NEP (or NEE). Weaker Rh trends than NPP trends resulted in a net uptake of carbon by boreal ecosystems during 1982-1998 (NEE trend of -2.9 gC m<sup>-2</sup> yr<sup>-2</sup>) (Lucht et al., 2002). Also on a global scale, ecosystem models simulate stronger NPP than Rh trends and thus an increasing global net land uptake in the last decades of the 20<sup>th</sup> century (Sitch et al., 2008, 2013). Therefore satellite observations of greening trends are indicative of positive trends in photosynthetic carbon uptake.

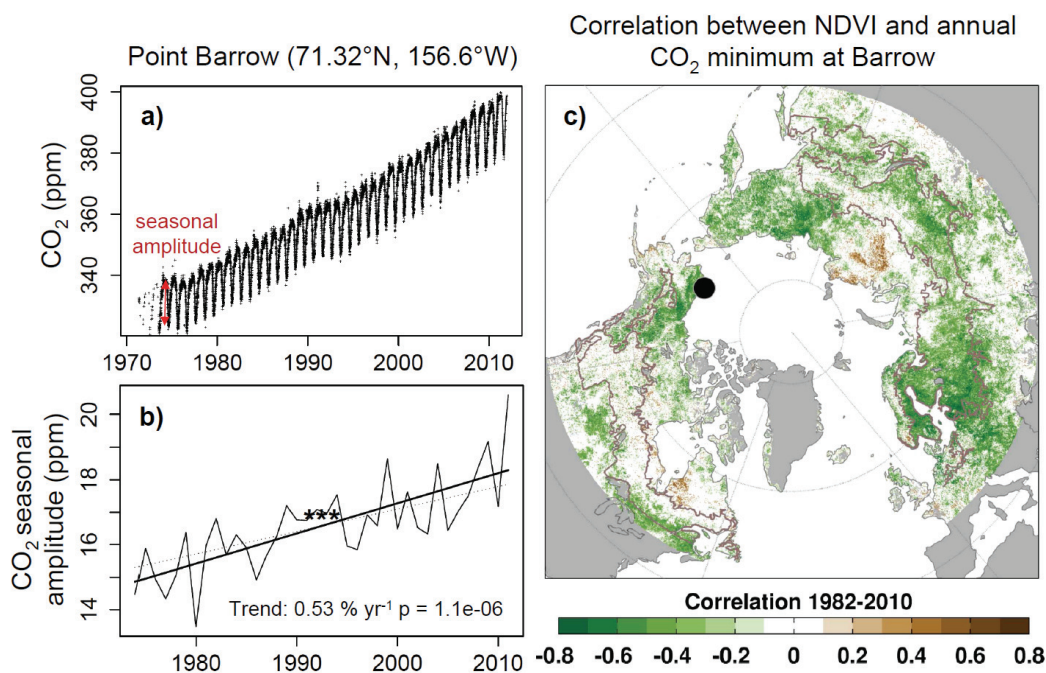


Figure 1.6: High-latitude greening and increasing amplitude of atmospheric CO<sub>2</sub>. (a) Time series of atmospheric CO<sub>2</sub> as observed at Point Barrow, Alaska and (b) annual seasonal amplitude of CO<sub>2</sub> and trend. Figures (a) and (b) are taken from manuscript 4, data from NOAA, (Conway et al., 1994). (c) Correlation between growing season-integrated NDVI (GIMMS3g dataset) and the seasonal minimum of de-trended CO<sub>2</sub> at Barrow (black dot) from 1982 to 2010. Only significant correlations are shown ( $p < 0.1$ ). Figure (c) is courtesy of Jonathan Barichivich (Barichivich et al., 2013).

Trends in the terrestrial net carbon balance are likely to affect atmospheric CO<sub>2</sub> concentrations (Le Quéré et al., 2009). Atmospheric monitoring demonstrates globally an increase in CO<sub>2</sub> because of fossil fuel emissions (Keeling et al., 1976). Additionally, it has been shown that the seasonal amplitude of atmospheric CO<sub>2</sub> increased over the last four decades of the 20<sup>th</sup> century especially at high latitude monitoring stations (Figure 1.6b) (Bacastow et al., 1985; Keeling et al., 1996). This CO<sub>2</sub> amplitude increase was also observed in the middle and upper troposphere by aircraft observations and is likely linked to changes in boreal and arctic ecosystems (Graven et al., 2013). Indeed, the increase of the CO<sub>2</sub> amplitude correlates well with greening trends in northern high latitudes (Figure 1.6c) (Barichivich et al., 2013; Myneni et al., 1997a). On the other hand, it has been hypothesized that the intensification of agriculture in temperate regions is an important control on the CO<sub>2</sub> amplitude trend (Gray et al., 2014; Zeng et al., 2014). However, these studies did not comprehensively account for other possible controlling factors and ignored especially the role of atmospheric transport. The co-occurrence of greening and CO<sub>2</sub> amplitude trends in high latitudes suggests a possible relationship. Nevertheless, despite of correlation analyses the effect of greening on the increase in the seasonal amplitude of atmospheric CO<sub>2</sub> has not yet been demonstrated and quantified.

The current role of the terrestrial land surface as carbon sink in the Earth system might change under future climate conditions but is highly uncertain. While some global climate/carbon-cycle models project a terrestrial carbon sink until 2100, others project a terrestrial carbon source (Friedlingstein et al., 2006, 2014). This spread among models is related to differences in the simulation of land use change emissions and to the sensitivity of the simulated terrestrial carbon cycle to rising atmospheric CO<sub>2</sub> (Friedlingstein et al., 2014). Especially global vegetation models diverge regarding their future projected vegetation residence time which defines if a model will project a carbon source or sink (Friend et al., 2014). Generally, global coupled climate/carbon-cycle models underestimate the turnover time of carbon in terrestrial ecosystems in comparison to data-oriented estimates which might be related to a too weak sensitivity of models to precipitation (Carvalho et al., 2014). The spread in simulated vegetation residence times among models is related to how demographic processes such as recruitment, mortality and vegetation composition are considered (Friend et al., 2014).

Satellite observations of vegetation greenness provide information about changes in vegetation phenology, type and vegetated area. Therefore these observations could partly help to improve some of these processes in ecosystem models that contribute to the uncertainties in the future development of the land carbon cycle feedback.

### 1.3 Remote sensing of greenness and phenology

#### 1.3.1 Vegetation indices and biophysical properties

The greenness of the vegetation is quantified from satellite observations by vegetation indices. A vegetation index is a combination of spectral bands from an optical sensor that contain information about spectral properties of vegetation (Huete, 2012). Green healthy vegetation absorbs radiation especially in the blue and red regions of the electromagnetic spectrum and reflects a large part of the incoming radiation in the near infra-red (NIR) regions (Figure 1.7a). On the other hand, soil has an increasing reflectance from the blue to the NIR region. These distinct differences between red and NIR absorption by vegetation and bare soil allow the formulation of the Normalized Difference Vegetation Index (NDVI) to quantify these differences (Tucker, 1979):

$$NDVI = \frac{NIR - red}{NIR + red} \quad (1.3)$$

NDVI ranges between -1 and 1 whereas vegetated surfaces usually have NDVI values > 0.2. Bare soil, snow and water have lower NDVI values.

The greenness of vegetation affects the carbon cycle through the absorption of radiation for photosynthesis. The absorbed photosynthetic active radiation (*APAR*) depends on the total incoming photosynthetic active radiation (*PAR*) and on the greenness of the vegetation:

$$APAR = FAPAR \times PAR \quad (1.4)$$

where *FAPAR* is the fraction of absorbed photosynthetic active radiation (Monteith, 1972). Moreover, *FAPAR* has a linear relationship to NDVI (Huete, 2012; Myneni and Williams, 1994) although this relationship is variable depending on site conditions, plant types or the development stage of the canopy (Fensholt et al., 2004; Gamon et al., 1995; Myneni et al., 1997b). The amount of leaves in a canopy affects NDVI (*FAPAR*). The amount of leaves is often quantified as leaf area index (LAI) which is the one-sided area of all leaves ( $A_{leaves}$ ) in relation to the ground area ( $A_{ground}$ ) (Bonan, 2008a, p.253):

$$LAI = A_{leaves} / A_{ground} \quad (1.5)$$

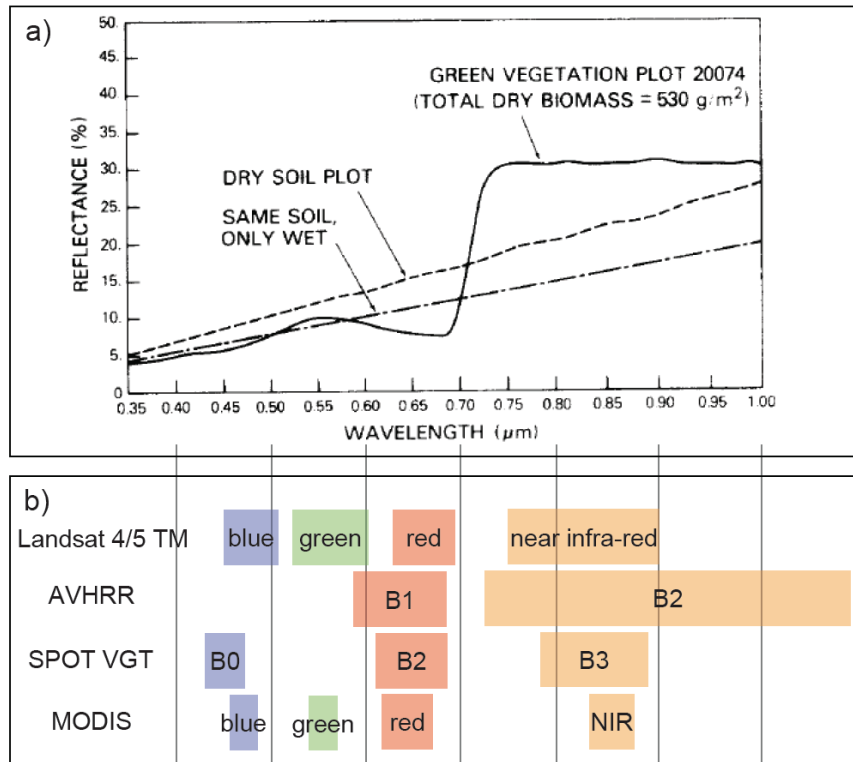


Figure 1.7: Reflectance of vegetation, soil and location of spectral bands from different satellite sensors.

(a) Spectral reflectance for green vegetation, dry soil, and wet soil in the visible and near infrared part of the electromagnetic spectrum (from Tucker, 1979). (b) Location of spectral bands from different satellite sensors that are used to retrieve long-term global datasets of vegetation greenness (Baret et al., 2007; Huete et al., 2002; Pinzon and Tucker, 2014; Richards and Jia, 2006, p.5).

If LAI becomes higher (e.g. during leaf development in spring) more PAR is absorbed by the canopy and thus FAPAR (NDVI) increases. Nevertheless, at high LAI values FAPAR (NDVI) increases only marginally or saturates because this relationship follows approximately a Lambert-Beer relation of light absorption (Monsi and Saeki, 1953, 2005):

$$FAPAR = 1 - e^{-k \times LAI} \quad (1.6)$$

where  $k$  is the light extinction coefficient which depends on solar zenith angle, leaf orientation, clumping and thus plant species (Chen et al., 2005; Monsi and Saeki, 1953). Nevertheless, this simple relationship of canopy radiation transfer assumes that light is completely absorbed and no reflection or scattering within the canopy occurs (Bonan, 2008a, p.255; Sellers, 1985). As FAPAR (NDVI) depends on the amount of leaves, it is

also closely related to the fractional area coverage of green vegetation which can be expressed as fractional vegetation cover (Carlson and Ripley, 1997) or foliar projective cover (FPC) (Specht, 1972, 1981). FPC is the fractional coverage of the ground by leaves (Huete, 2012). Other vegetation indices that account for soil reflectance, that have a higher sensitivity to photosynthetic light absorption, or that are sensitive to water availability exist and are reviewed in Huete et al. (2012). However, such vegetation indices are not readily available from long-term (30 years) harmonized satellite datasets and could therefore not be used in this thesis.

### **1.3.2 Observational uncertainties from sensors and datasets**

Vegetation index time series are prone to uncertainties. One major source of uncertainty originates from the availability of several datasets that differ by the used satellite sensor with associated spectral properties and retrieval algorithms. Different “families” of satellite sensors have been used to derive long-term time series of vegetation greenness (Figure 1.7b), e.g.:

- The Landsat program provides the longest record of relative high spatial resolution satellite images since 1972 (NASA, 2015; Richards and Jia, 2006, p.393). However, Landsat imagery has only been used in regional studies to assess long-term changes in vegetation greenness (Baird and Verbyla, 2012; Fraser et al., 2011; Neigh et al., 2008). This is due to the initially high cost for imagery, the sparse temporal sampling, the high spatial resolution and thus high requirements for data storage, and the relatively high efforts that are needed to process Landsat data (atmospheric correction, cloud removal, inter-sensor calibration).
- AVHRR (Advanced Very High Resolution Radiometer) instruments on board NOAA (National Oceanic and Atmospheric Administration) satellites provide the second longest and most used records of greenness time series starting in July 1981 (Pinzon and Tucker, 2014; Tucker et al., 2005). AVHRR sensors have a very coarse spatial (4 \* 4 km) and spectral resolution (Figure 1.7b). Nevertheless, these sensors provide a daily coverage of the Earth which is the crucial advantage of using these sensors to derive time series of vegetation greenness.
- MODIS (Moderate Resolution Imaging Spectroradiometer) instruments on Terra and Aqua satellites were launched 1999 and 2002, respectively. The spectral

properties of the MODIS sensor and its medium spatial resolution (250 to 1000 m depending on spectral band) makes it suitable for large scale assessments of vegetation greenness dynamics (Huete et al., 2002). Therefore MODIS-derived greenness time series are often used as a benchmark for the longer AVHRR-derived time series (Fensholt et al., 2009; Fensholt and Proud, 2012; Tucker et al., 2005).

- Vegetation (VGT) instruments onboard the SPOT 4 and 5 (Satellite Pour l'Observation de la Terre) and Proba-V satellites were launched 1998, 2002 and 2013, respectively. The spectral (Figure 1.7b) and spatial resolution (~300 m) and the approximately daily return interval of this sensor has a similar suitability for large-scale vegetation observation as MODIS (Baret et al., 2007, 2013).

Other vegetation index datasets have been derived from the SeaWiFS (Sea-viewing Wide Field-of-view Sensor, 1997-2011) and MERIS (Medium Resolution Imaging Spectrometer, 2002-2012) sensors (Gobron et al., 2006, 2008).

The different orbital and spectral properties of these satellites and sensors result in different estimates of vegetation indices. Moreover satellite sensors are affected by orbital drift (Kaufmann et al., 2000) and from sensor degradation (Wang et al., 2012) which result in a changing sensitivity of the sensors to the received signal. Additionally, high sun zenith angles, clouds, aerosols, or snow cover can distort observations and frequently result in an underestimation of NDVI (Holben, 1986). Therefore, satellite observations need to be calibrated, multiple sensor observations need to be harmonized and measurements need to be corrected for distortions to ensure consistent long-term time series of vegetation greenness. A major sensor calibration and across-sensor harmonization effort was achieved by the GIMMS (Global Inventory Monitoring and Modelling Study) group that created a global NDVI dataset with 8 km spatial resolution starting in July 1981 based on data from AVHRR sensors (Tucker et al., 2005). This NDVI dataset has recently been newly calibrated and updated to the present (GIMMS3g, 3<sup>rd</sup> generation, Pinzon and Tucker (2014)). FAPAR and LAI datasets were derived from the GIMMS3g NDVI dataset as well (Zhu et al., 2013). Vegetation indices were also retrieved from other sensors using different retrieval algorithms (Baret et al., 2007, 2013; Gobron et al., 2006; Gobron and Verstraete, 2009; Knyazikhin et al., 1999; Pinty et al., 2007; Plummer et al., 2007). Nevertheless, the

GIMMS NDVI dataset has been most widely used for the analysis of trends in land surface phenology and greenness because of the longer time coverage.

Despite harmonization efforts, vegetation index datasets from different sensors or processing algorithms still show remarkable differences. For example, FAPAR datasets differ in mean values and seasonality across different ecosystems (D'Odorico et al., 2014; McCallum et al., 2010). NDVI time series from different datasets differ in the timing of spring onset or start of growing season (Atzberger et al., 2013; White et al., 2014). Moreover, the extent and spatial distribution of greening and browning trends in boreal forests does not agree among datasets which led to doubts about the reliability, especially of browning trends (Alcaraz-Segura et al., 2010). Although datasets from other satellites generally confirmed the existence of browning trends in boreal forests of North America (Beck and Goetz, 2011; Parent and Verbyla, 2010), the new GIMMS3g version shows less extensive browning trends than its predecessor (Jiang et al., 2013). Greening occurs more in deciduous boreal forests and browning occurs more in evergreen boreal forests (Beck and Goetz, 2011). On the other hand, datasets generally agree less on the scale and direction of these trends in evergreen forests (McCallum et al., 2010). Differences in the temporal dynamics between satellite datasets were found across the world (Fensholt and Proud, 2012) such as the tundra (McCallum et al., 2010), temperate grasslands (D'Odorico et al., 2014; Scheftic et al., 2014), tropical savannahs like the Sahel (Fensholt et al., 2009), and tropical forests (Morton et al., 2014). Consequently, studies of land surface phenology and greenness should consider such uncertainties from datasets to robustly assess ecological changes.

### **1.3.3 Observational uncertainties from time series analysis methods**

An additional source of uncertainty in the analysis of satellite-observed vegetation greenness dynamics arises from time series analysis methods. The availability and specific advantages of different time series analysis methods affects the analysis of land surface phenology as well as of long-term greenness trends.

Satellite-derived vegetation index time series require pre-processing in order to assess land surface phenology (de Beurs and Henebry, 2010b). Some datasets have gaps because of the removal of observations under cloudy conditions or snow cover. These gaps, noise, and the coarse temporal resolution necessitate interpolation and

smoothing of time series to extract the main seasonal features. Several methods have been used for interpolation and smoothing of vegetation index time series (Beck et al., 2006; Delbart et al., 2006; Jonsson and Eklundh, 2002; Verbesselt et al., 2010b). However, the quality of these methods differs and it is difficult to identify a best method (Geng et al., 2014). The performance of methods decreases under higher levels of noise and is dependent on the length of gaps, seasonality, vegetation type and sensor characteristics (Geng et al., 2014; Hird and McDermid, 2009; Kandasamy et al., 2013; Musial et al., 2011). The varying performance of interpolation and smoothing methods contributes to uncertainties in the estimation phenology metrics from vegetation index time series. Metrics of land surface phenology are for example the start and end of the growing season (SOS or EOS) which define the length of the growing season (LOS). Nevertheless, no generally accepted definition or detection methods exist for these metrics which introduces remarkable uncertainties (Beck et al., 2006; de Beurs and Henebry, 2010b; Elmore et al., 2012; Jonsson and Eklundh, 2002; Karlsen et al., 2006; Tateishi and Ebata, 2004; White et al., 2014, 2009). For example, methods can differ in up to two months in estimated SOS (White et al., 2014, 2009). Consequently, studies of land surface phenology should consider different time series interpolation and smoothing methods, and different methods for the extraction of phenology metrics for a robust assessment of land surface phenology dynamics.

Different methods exist to quantify trends in land surface phenology or greenness. A frequently used method is to estimate the trend as the slope of a linear regression. However, the use of linear regressions violates statistical assumptions such as the independence of observations due to temporal autocorrelation or homogeneity (de Beurs and Henebry, 2004b). Alternative approaches consider temporal autocorrelation structures (Goetz et al., 2005) or non-parametric trend tests, like the Mann-Kendall test (Kendall, 1975; Mann, 1945), to circumvent the limitations of regression analysis (de Beurs and Henebry, 2004b, 2010a). Although different trend methods agree on the overall patterns of the major greening and browning trends, they reveal clear differences in regions with weak trends (de Jong et al., 2011a) which suggests that the performance of trend methods depends on the trend magnitude. Additionally, trends in NDVI time series are often non-monotonic, i.e. they might change because of climate variability, disturbances or land cover change (de Beurs and Henebry, 2004a; de Jong et



al., 2011b, 2013b). Therefore methods were developed that consider abrupt or gradual changes in trends (Bai and Perron, 2003; Mudelsee, 2009; Verbesselt et al., 2010a, 2010b). De Jong et al. (2013b) found trend changes in the GIMMS3g NDVI dataset globally, like changes from initial greening to browning. However, the performance of different trend methods and the reliability of trend changes have not yet been investigated, especially in arctic and boreal ecosystems, which are the most prominent examples of such trends.

## **1.4 Model-data integration for greenness and phenology**

Global ecosystem models have previously been used to assess the importance of different environmental controls over greening trends or to quantify their effects on terrestrial net primary production (Lucht et al., 2002; Mao et al., 2012, 2013; Nemani et al., 2003; Piao et al., 2006). This chapter provides an overview of global ecosystem models; provides an overview about how current state-of-the-art model reproduce observed phenology and long-term greenness dynamics; introduces concepts of model-data integration; and finally provides examples of how phenology models can be improved through model-data integration.

### **1.4.1 Dimensions and groups of global ecosystem models**

Global ecosystem models are simplified representations of the terrestrial biosphere that serve different purposes, such as the assessment of the distribution of vegetation types or the quantification of carbon and water fluxes in ecosystems. In a broader sense, classification schemes like the climate classification of Köppen (1884) are models of the distribution of global vegetation. However, in a narrower sense global ecosystem models are codified formulations which are used to quantify states and responses of terrestrial ecosystems to environmental conditions (Prentice et al., 2007). The comprehensiveness of the biosphere representation depends mostly on the historical origins of a model and the main area of application. Consequently global ecosystem models can be grouped according to different dimensions (Figure 1.8). Models can be grouped according to their main application domain (diagnostic to prognostic models), according to the level of physical process representation (empirical to

physical/mechanistic), and according to how comprehensively they account for ecosystem structure and processes. However transitions between model groups are smooth. Diagnostic models are used to assess ecosystems in the present or in the recent past. Prognostic models are applied to assess ecosystem responses in the future or in the far past where ecosystem states cannot be constrained by observations and thus need to be fully simulated.

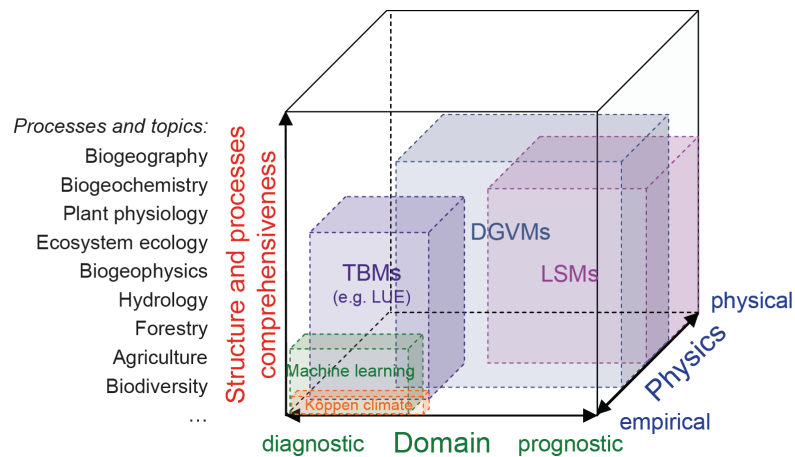


Figure 1.8: Dimensions of global ecosystem models.

Empirical models depend on observational data. For example, machine learning approaches have been used to estimate spatial-temporal patterns of ecosystem carbon and water fluxes based on ecosystem-level and satellite observations (Papale and Valentini, 2003; Reichstein et al., 2007) and, more recently to quantify ecosystem carbon and water fluxes globally (Beer et al., 2010; Jung et al., 2011). Although machine learning approaches are simplified representations of biosphere processes, they are not generally seen as ecosystem models in a narrow sense but rather as data-oriented estimates because they do not depend on assumptions about physical relations in ecosystems.

Global ecosystem models are sometimes distinguished between terrestrial biogeochemical models (TBM) and dynamic global vegetation models (DGVM). TBMs simulate the carbon and water cycle of terrestrial ecosystems by assuming a fixed ecosystem composition. DGVMs simulate additionally changes in ecosystem composition and structure (Bonan, 2008a, p.379; Prentice et al., 2007). Historically, the primary aims of TBMs and DGVMs was to simulate net primary production and the distribution of global vegetation, respectively (Prentice et al., 2007). Both groups of

models represent ecosystem processes based on a wide range of empirical to physical formulations. For example, a group of TBMs are light-use efficiency models (LUE, also radiation-use efficiency, RUE) that rely on an empirical relation to estimate NPP or GPP from FAPAR (Monteith, 1972; Potter et al., 1999; Running et al., 2000):

$$GPP = FAPAR \times PAR \times LUE_{max} \times f(T, W, \dots) \quad (1.7)$$

where  $LUE_{max}$  is the maximum light-use efficiency, i.e. the maximum amount of assimilated carbon per unit light, and  $f(T, W, \dots)$  are scalar functions that account for environmental stresses such as temperature ( $T$ ) and water stress ( $W$ ). The definition of these stress functions differs among implementations of LUE-models which causes large differences in simulated GPP (Cai et al., 2014). LUE-models are also diagnostic models because they rely on (satellite) datasets of FAPAR or NDVI and have therefore no prognostic capabilities. Other TBMs which are prognostic simulate photosynthesis based on mechanistic formulations (Knorr, 2000; Melillo et al., 1993).

Another group of global vegetation models was originally developed within global climate and weather forecast models to account for the exchange of energy, heat and momentum between the land surface and atmosphere; namely land surface models (LSMs) (Sellers et al., 1997). Later, LSMs were developed further to fully account for terrestrial carbon and water cycles and thus developed into full global ecosystem models (Bonan, 2008a, p.396; Sellers et al., 1997). However, prognostic TBMs and LSMs had a limited capability to be used for long-term projections of climate impacts on vegetation because they did not account for the effects of climate change on changes in vegetation structure and composition.

Therefore DGVMs were developed to comprehensively account for the impacts of climate change on vegetation dynamics, i.e. plant growth, competition, mortality and ecosystem structure (Cramer et al., 2001). DGVMs represent vegetation as a set of plant functional types (PFTs). PFTs are often defined based on biome (tropical, temperate, boreal), plant form (tree, grass), leaf type (needle-leaved, broad-leaved) and phenology (evergreen, deciduous) (Prentice et al., 2007; Sitch et al., 2003). The development of TBMs, LSMs and DGVMs converges nowadays by increasing the comprehensiveness of represented ecosystem structural and process components in all of these models. Recent developments in global ecosystem models are for example the incorporation of human-induced land use change and agriculture (Bondeau et al., 2007; Gervois et al.,

2004; Kucharik, 2003), disturbance as fire (Kloster et al., 2010; Thonicke et al., 2001, 2010; Venevsky et al., 2002), or the consideration of biodiversity by replacing fixed PFT-parameterizations with variable plant traits (Bodegom et al., 2014; Pavlick et al., 2013). DGVMs have been adapted to atmospheric models and thus increasingly replace classical LSMs in coupled biosphere-atmosphere models (Bonan et al., 2003; Foley et al., 2000). Based on this convergence in model development barriers between TBMs, LSMs and DGVMs disappear and such a distinction becomes obsolete.

It is evident to use an ecosystem model (DGVM) that considers vegetation dynamics to assess environmental controls on vegetation greenness and phenology. Although most prognostic ecosystem models simulate phenology and vegetation greenness (FAPAR), only a DGVM is suitable for such an application because greening and browning trends are related to changes in vegetated area and vegetation composition. Vegetation composition in DGVMs depends thereby on bioclimatic limits, net primary production, carbon allocation, competition, mortality, and disturbance (Cramer et al., 2001; Prentice et al., 2007; Sitch et al., 2003). For example, the LPJ (Lund-Potsdam-Jena) DGVM simulates the allocation of the annually assimilated carbon (NPP) to leaves, wood and roots (Sitch et al., 2003). Increasing leaf carbon ( $C_{leaf}$ ) results in an increasing LAI of individual trees:

$$LAI_{ind} = \frac{C_{leaf} \times SLA}{CA} \quad (1.8)$$

where  $SLA$  and  $CA$  are the specific leaf area ( $m^2 gC^{-1}$ ) and crown area, respectively (Sitch et al., 2003). LAI is then converted to FPC and finally to FAPAR by using the Lambert-Beer equation (eq. 1.6). Therefore DGVMs such as LPJ can be used diagnostically (as in Lucht et al. (2002)) to assess controlling factors for land surface phenology and greenness dynamics and to assess effects on the global carbon cycle.

#### 1.4.2 Evaluation of phenology and greenness in global ecosystem models

Recent studies have shown that global ecosystem models have a poor performance in representing observed greenness dynamics. For example, most ecosystem models simulate a too early start of the growing season in deciduous forests of North America in comparison to ground measurements of LAI (Richardson et al., 2012). The same study also shows that the inter-annual variability of the start of the growing season is

poorly reproduced by ecosystem models. These misrepresentations of phenology also resulted in an overestimation of GPP in spring and autumn (Richardson et al., 2012). Although ground observations are probably less uncertain than satellite observations, they provide only a limited insight in the large spatial performance of ecosystem models. A global evaluation of FAPAR simulations from the ORCHIDEE model against MODIS NDVI time series was done by Maignan et al. (2011). They found high correlations ( $> 0.8$ ) in most deciduous ecosystems (Europe, North America, Savannas), weak correlations ( $\sim 0.6$ ) in most boreal forests and the tundra, and low correlations for tropical forests and some drylands. Although such a correlation analysis between observed NDVI and modelled FAPAR gives a general notion on problems of modelled phenology, it is too simplistic because the NDVI-FAPAR relationship depends on vegetation type and phenological status. Therefore, differences in correlations between regions might not only reflect a poor model performance but also differences in the NDVI-FAPAR relationship. Thus, a model evaluation of simulated FAPAR (or LAI) against satellite-derived FAPAR (or LAI) estimates would be a more direct approach to assess spatial patterns of model performance. Thereby FAPAR estimates are closer to the actual spectral measurements than LAI estimates (Kelley et al., 2013).

A set of DGVMs was evaluated against the GIMMS LAI dataset over the Northern Hemisphere for the period 1986-2005 to assess the model performance for vegetation phenology and long-term greenness trends (Murray-Tortarolo et al., 2013). All models overestimated LAI over boreal forests and some models overestimated high-latitude greening trends whereas other models did not simulate greening at all (Murray-Tortarolo et al., 2013). Murray-Tortarolo et al. (2013) also evaluated simulated start and end of growing season dates but did not consider the above-mentioned observational uncertainties such as from different satellite datasets or phenology methods and therefore provides no insight in the reliability of these model evaluation results. Despite these methodological weaknesses, all studies (Anav et al., 2013; Kelley et al., 2013; Murray-Tortarolo et al., 2013; Richardson et al., 2012) demonstrate that global ecosystem model need to be improved to better represent seasonal, inter-annual and long-term greenness dynamics especially through a better representation of phenology.

### 1.4.3 Concepts of model-data integration

Observations or “data” are necessary during several steps of model development (Figure 1.9). Observations are especially needed during model characterization as consistency checks, during the estimation of model parameters, and for evaluation of model results (Williams et al., 2009). This concept is known as model-data fusion (MDF), model-data synthesis (MDS), data assimilation (DA), or model-data integration (MDI) (Keenan et al., 2011; Mathieu and O’Neill, 2008; Raupach et al., 2005; Wang et al., 2009). MDI is the combination of information from observations with information from models to evaluate models, to estimate model parameters or model states in order to quantify and reduce the uncertainty of the model structure, parameters, states or predictions (Liu and Gupta, 2007; Raupach et al., 2005).

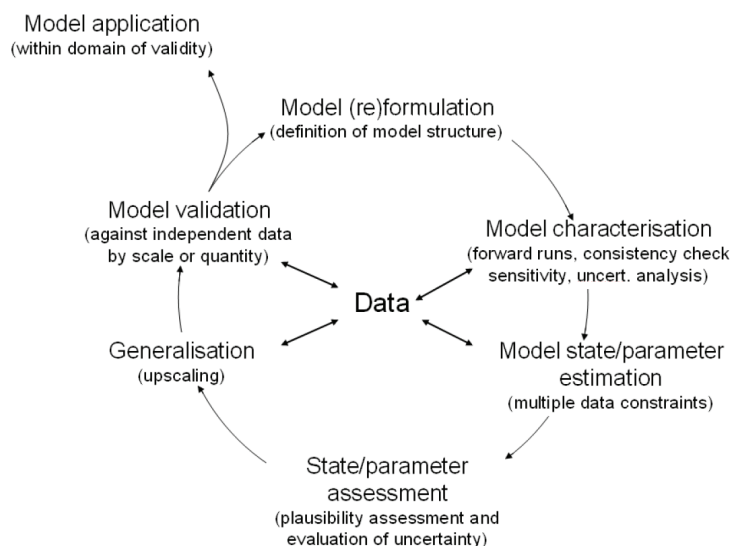


Figure 1.9: Use of data in model development and model-data integration.

The figure is taken from Williams et al. (2009) and was described in the context of FLUXNET eddy covariance data. The same model-data integration process can be also applied to satellite-derived greenness time series and prognostic ecosystem models such as DGVMs.

In a narrow sense, MDI is restricted to parameter and state estimation (Liu and Gupta, 2007; Raupach et al., 2005; Wang et al., 2009). In the following MDI is explained in the context of parameter optimization. Parameters are time-invariant properties of the model that usually define the sensitivity of a response to a forcing variable (Liu and Gupta, 2007). Generally, a MDI framework for parameter optimization consists of several components (Figure 1.10). The core of this framework is an ecosystem model that simulates an output as response to an input and model parameters (Liu and

Gupta, 2007). The model output is compared to an appropriate observed variable in a cost function. The cost function quantifies the mismatch between model and observation. An optimization method is used to reduce the cost by automatically adjusting model parameters. Cost functions are often defined based on the sum of squared errors (Keenan et al., 2011; Liu and Gupta, 2007; Raupach et al., 2005; Wang et al., 2009):

$$J = \sum_{t=1}^n \left( \frac{m_t(x) - o_t}{\sigma_t} \right)^2 \quad (1.9)$$

where  $m$  and  $o$  are the modelled and observed quantities at time  $t$ , respectively.  $\sigma$  is the uncertainty of observations and  $x$  is a vector of model parameters that needs to be adjusted in order to reduce  $J$ . In Bayesian optimization approaches the uncertainty of models parameters is considered in the cost function as well (Liu and Gupta, 2007). Cost functions can be also defined based on other model performance metrics (Janssen and Heuberger, 1995; Krause et al., 2005) or based on combined metrics (Gupta et al., 2009).

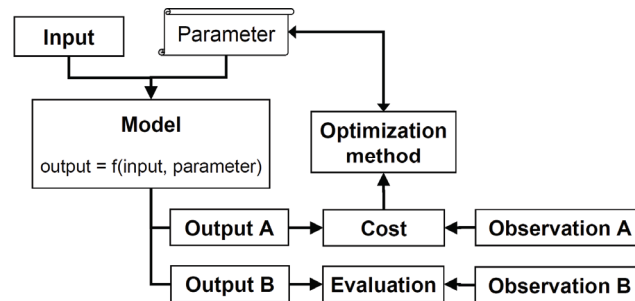


Figure 1.10: Schematic diagram of the structure of a model-data integration framework.

The reduction of the cost between model and observation is achieved by applying optimization techniques (Trudinger et al., 2007). In general, local and global search optimization techniques are distinguished (Wang et al., 2009). Local search (or gradient-based) methods such as the Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963) or BFGS (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970) algorithms use the first or second derivative of the model to find an optimum. Local search methods are computationally fast but are prone to result in a local optimum dependent on the initial parameter values. Alternatively, global search methods such as genetic optimization algorithms (Mitchell, 1998) or Markov Chain Monte Carlo

approaches (Metropolis et al., 1953) explore the full parameter space, thus more likely converging towards the global optimum parameter set and allow a non-parametric characterization of posterior parameter uncertainties. However, it has been shown that the variability in estimated parameters terrestrial biogeochemical models is more affected by the definition of the cost function than by the chosen optimization algorithm (Trudinger et al., 2007). Cost functions can have several local minima, a parameter might be insensitive to the cost function, or parameters are correlated, i.e. a reduced cost can be achieved with different combinations of model parameters; referred to as equifinality (Wang et al., 2009). Equifinality can be avoided by reducing the number of considered parameters (and thus changing the model structure); by using narrower prior estimates for parameter ranges; or by using additional observations that might help to distinguish correlated parameters (Carvalhais et al., 2010; Wang et al., 2009).

Some general rules of good practice should be considered within parameter estimation in ecosystem models (Keenan et al., 2011): 1) Observational uncertainties should be quantified and included within parameter optimization; 2) Alternative model structures should be explored and tested; 3) Optimization should be done against multiple data streams; 4) The MDI system should be tested by using synthetic data from a forward model run; 5) The model should be evaluated against independent observations; 6) Posterior uncertainties of model parameters and predictions should be quantified. These rules will help to efficiently reduce parameter uncertainties, to avoid equifinality, and to avoid over-fitting of the model to a specific data set.

Satellite-derived vegetation indices are useful data sets in model-data integration because they contain information about phenology and vegetation dynamics (Prentice et al., 2007). Satellite observation of vegetation indices can be used to force diagnostic and empirical ecosystem models (Running et al., 2000), to estimate states of ecosystem models (Quaife et al., 2008), or to estimate model parameters of mechanistic ecosystem models (Dorigo et al., 2007). However, observational uncertainties of satellite-derived vegetation index data sets need to be considered in model-data integration (Raupach et al., 2005). In summary, the use of satellite-derived vegetation index time series within a model-data integration framework can potentially help to overcome limitations of



current global ecosystem models in representing observed phenology and long-term greenness dynamics.

#### 1.4.4 Improved phenology modelling through model-data integration

Phenology in ecosystem models can be improved by either reformulating the model structure and/or by optimizing model parameters (Figure 1.9). Migliavacca et al. (2012) have shown that model structural uncertainty is of higher importance than parameter uncertainty for the uncertainty of future climate change scenarios of phenology. This finding requires a careful characterisation and evaluation of current phenology model structures. Many prognostic ecosystem models describe the seasonal leaf development dependent on growing degree days (GDD), i.e. leaves emerge and develop as heat accumulates during spring (Richardson et al., 2012). Although this concept might be appropriate for temperate deciduous trees, it fails for temperate grass ecosystems that have an additional strong control by water availability (Kramer et al., 2000; Liu et al., 2013; Yuan et al., 2007). Therefore it might be necessary to test alternative phenology models to current model structures that consider alternative or additional controlling factors for phenology (Migliavacca et al., 2008, 2012). A modelling approach to consider different factors for phenology is the Growing Season Index (GSI) (Jolly et al., 2005). GSI considers different climate limits on the development of foliage and is an index between 0 (no foliage, or inactive) and 1 (full developed foliage). The original GSI is calculated from three multiplicative functions that account for effects of minimum temperature ( $T_{\min}$ ), vapour pressure deficit (VPD) and day length (Jolly et al., 2005):

$$GSI = f(T_{\min}) \times f(VPD) \times f(daylength) \quad (1.10)$$

The three functions are stepwise linear functions that are defined based on certain threshold parameters. However these parameters are empirical and need to be estimated. For example, it has been shown that the original parameters of the GSI model needed to be optimized to accurately reproduce the seasonality of larch trees in the Alps (Migliavacca et al., 2008). Moreover the GSI approach is valuable because it allows easy consideration of other potentially influencing factors on phenology by simply multiplying additional controlling functions. For example, the phenology of subalpine grasslands was best modelled using the GSI approach when the VPD-

function was replaced by a function depending on soil water content and by additionally accounting for snow cover (Migliavacca et al., 2011). The GSI functions provide an estimation of the climatic constraints and therefore can be used to map the importance of climatic factors on phenology (Jolly et al., 2005; Migliavacca et al., 2008; Stöckli et al., 2008b). Nevertheless, adding a new functional relation to the GSI approach requires estimation of the empirical parameters. Parameters of GSI were optimized against MODIS LAI and FAPAR products globally which demonstrated the possibility of successfully applying GSI for a variety of biomes (Stöckli et al., 2011). Nevertheless, GSI is a stand-alone phenology model and not linked to a full ecosystem model.

The optimization of a phenology model within a full global ecosystem model is even more challenging because it is additionally necessary to constrain processes like photosynthesis. For example, phenology- and photosynthesis-related parameters have been for example optimized in the BETHY model against satellite FAPAR data and atmospheric CO<sub>2</sub> measurements (Kaminski et al., 2012), and in the ORCHIDEE model against site eddy covariance measurements of ecosystem carbon and water fluxes (Kuppel et al., 2012, 2014). These model optimization studies resulted in improved representations of phenology and in a reduced uncertainty of simulated regional and global carbon and water fluxes. However, these model optimizations studies were only done based on relatively short observational time series (~ 5 years). It is therefore not clear if these efforts also result in an improved representation of long-term greenness or carbon cycle dynamics.

In summary, improvements of global ecosystem models through reformulations of model structures and parameter optimization need to be performed with respect to long observational time series ( $\geq 30$  years). Such long time series allow assessment of the ability of models to reproduce recent observed phenology, greenness, and carbon cycle dynamics like greening or browning trends, and positive trends in the seasonal CO<sub>2</sub> amplitude. By considering observational uncertainties the model could be applied to identify environmental controls on these recent dynamics in the Earth system.

## 1.5 Research questions

As mentioned above, the overall aim of this thesis is to *better understand recent dynamics of land surface phenology and greenness in the Earth system* by combining satellite observations with global vegetation models. *Land surface phenology and greenness* refers to the timing of seasonal changes in vegetation (phenology) and to instantaneous or temporally averaged values (greenness) of vegetation index (NDVI or FAPAR) time series as derived from satellite observations or model simulations of the land surface. *Recent dynamics* involve multiple time scales like seasonal changes, inter-annual variability and trends between 1982 (1970 in chapter 5) and 2011. *Understanding* means 1) objectively quantifying these dynamics from observations while accounting for uncertainties; 2) modelling these dynamics with process-based ecosystem models while contrasting models with observations; 3) explaining the controlling factors for these dynamics; and 4) quantifying their effects (especially of greening trends) on the terrestrial carbon cycle.

### 1.5.1 How do observational uncertainties contribute to the quantification of dynamics in land surface phenology and greenness?

The quantification of dynamics in land surface phenology and greenness is affected by several observational uncertainties that originate from different satellite datasets and time series processing. Trend analysis methods differ in how they account for seasonality in time series. Additionally, some recently developed trend methods account for non-monotonic changes that can be caused by abrupt changes like disturbances and thus can be identified as breakpoints in time series (Verbesselt et al., 2010a, 2010b). However, the reliability of such trend change detection methods (i.e. detecting breakpoints in trends) has not yet been assessed in boreal and arctic ecosystems. *This thesis aims to assess the performance of different trend change detection methods, and the reliability of estimated trends and breakpoints in boreal and arctic ecosystems* (chapter 2, Forkel et al., 2013). The performance of these methods was evaluated by simulating surrogate time series based on real time series properties of the GIMMS3g NDVI dataset. Furthermore, methods were applied to the GIMMS3g NDVI dataset over Alaska. Alaska is a suitable region to test trend change detection methods because

greening trends have been previously observed in the Alaskan tundra as well as browning trends in the boreal forests (Baird and Verbyla, 2012; Beck et al., 2011; Parent and Verbyla, 2010). Additionally, boreal forests in Alaska are often and severely affected by wild fires (Kasischke et al., 2010). Thus, the availability of long-term observation of fires can be used to assess the reliability of estimated breakpoints.

Time series interpolation, smoothing and phenology extraction methods, as well as different satellite datasets can give remarkably different results of phenology dynamics (de Beurs and Henebry, 2010b; White et al., 2014, 2009). The identification of a phenology method with an optimal performance depends on the availability of ground observations and is therefore not possible for global scale applications. *This thesis aims to quantify the observational uncertainty from different satellite datasets and phenology detection methods in global patterns and dynamics of land surface phenology and greenness* (chapter 4 and supplement in chapter 8). The quantification of these uncertainties allows a more robust evaluation of model simulations.

### **1.5.2 How can DGVMs be improved to better represent phenology and greenness dynamics?**

It has been shown that ecosystem models need an improved representation of phenology and long-term greenness dynamics (Murray-Tortarolo et al., 2013; Richardson et al., 2012). Phenology models have previously been improved by reformulating model structures or by optimizing model parameters (Kaminski et al., 2012; Kuppel et al., 2012, 2014; Migliavacca et al., 2012). Nevertheless, such model improvements have not yet been done with respect to long-term (30 years) observations of greenness dynamics. *This thesis aims to improve the representation of phenology in the LPJmL DGVM and to assess the impact of these model improvements on long-term greenness dynamics and carbon cycle simulations* (chapter 3 and supplement in chapter 7, Forkel et al., 2014). Thereby it is of special interest if model optimization at seasonal but long-term time series results also in an improvement regarding inter-annual variability and trends. The LPJmL DGVM (Bondeau et al., 2007; Sitch et al., 2003) has been selected as the central ecosystem model of this thesis because it has been previously used to explain greening trends (Lucht et al., 2002) and because it considers several processes that might be of importance for explaining trends in boreal and arctic

ecosystems like vegetation dynamics, fire (Thonicke et al., 2010) and permafrost (Schaphoff et al., 2013). Similar to other DGVMs, it has been shown that the phenology of LPJ requires improvements (Kelley et al., 2013; Mahecha et al., 2010b; Murray-Tortarolo et al., 2013; Richardson et al., 2012). An improved version of the LPJmL model that accurately reproduces observations of land surface phenology and greenness dynamics can be potentially applied to identify controlling factors for these trends (chapter 4) and furthermore might be useful in assessing the role of greening in recent carbon cycle trends (chapter 5).

### **1.5.3 What are the controls on land surface phenology and greenness on multiple time scales?**

Greening and phenology trends in high-latitude ecosystem have usually been related to warming temperatures (Lucht et al., 2002; Wang et al., 2011; Xu et al., 2013). A large proportion of the temporal dynamic of vegetation greenness is related to the seasonal development and senescence of leaves. Although leaf phenology is usually explained by temperature in temperate forests (Keenan et al., 2014), temperature control might be an insufficiently or an inaccurately represented controlling factor as most ecosystem models show significant limitations in reproducing the observed phenology (Richardson et al., 2012). *This thesis aims at a comprehensive assessment of the controls on vegetation greenness phenology by considering effects of light and water availability and their spatial-temporal importance* (chapter 3 and 7, Forkel et al., 2014). Furthermore, if other factors than temperature might be of importance for phenology, it is also necessary to assess their importance on inter-annual variability and trends in land surface phenology and greenness. For example, it has been suggested that fire (Goetz et al., 2005), drought (Buermann et al., 2014) and seasonal changes in snow and water availability (Barichivich et al., 2014) contribute to long-term greenness dynamics in boreal forests. *This thesis aims to quantify the regional importance of seasonal controlling factors like temperature, short-wave radiation, and water availability, and of factors like CO<sub>2</sub> fertilization, fire, permafrost, and land cover dynamics on the inter-annual variability and trends in land surface phenology and greenness* (chapter 4 and 8). Such a quantification of regional controls could potentially clarify the divergent continental patterns of more greening in boreal Eurasia than in boreal North America (Bi et al., 2013). These research

aims is addressed by performing factorial model experiments with LPJmL using an improved phenology module.

#### **1.5.4 What is the contribution of high latitude greening on the increasing amplitude of atmospheric CO<sub>2</sub>?**

The increasing seasonal amplitude of atmospheric CO<sub>2</sub> (Graven et al., 2013; Keeling et al., 1996) is concordant with the greening of boreal and arctic ecosystems (Barichivich et al., 2013; Myneni et al., 1997a). On the other hand, the intensification of agriculture has been suggested as an additional important controlling factor but without comprehensively accounting for other factors (Gray et al., 2014; Zeng et al., 2014). *This thesis aims to identify the controlling factors for the increase of the seasonal CO<sub>2</sub> amplitude by comprehensively considering other factors such as atmospheric transport* (chapter 5 and extended data in chapter 9). Specifically, the role of greening trends and changes in photosynthetic carbon uptake in boreal and arctic ecosystems will be quantified. To answer this research question, the LPJmL DGVM is coupled to the TM3 atmospheric transport model (Heimann et al., 1989) in order to simulate temporal dynamics of atmospheric CO<sub>2</sub>. The LPJmL model is evaluated against several independent data sets to analyse the plausibility of simulations of the recent carbon cycle. These independent datasets cover gross primary production (Jung et al., 2011), biomass (Saatchi et al., 2011; Thurner et al., 2014) (both in chapter 3), net biome productivity (Rödenbeck et al., 2003), and atmospheric CO<sub>2</sub> from ground observations (chapter 5).



## 2 Trend change detection in NDVI time series: effects of inter-annual variability and methodology

---

This chapter was originally published as:

Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M., Neigh, C. and Reichstein, M.: Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology, *Remote Sensing*, 5(5), 2113–2144, doi:10.3390/rs5052113, 2013.

*Remote Sens.* **2013**, *5*, 2113–2144; doi:10.3390/rs5052113

OPEN ACCESS

*Remote Sensing*

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

*Article*

### **Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology**

**Matthias Forkel <sup>1,\*</sup>, Nuno Carvalhais <sup>1,2</sup>, Jan Verbesselt <sup>3</sup>, Miguel D. Mahecha <sup>1</sup>, Christopher S.R. Neigh <sup>4</sup> and Markus Reichstein <sup>1</sup>**

<sup>1</sup> Department for Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany; E-Mails: ncarval@bgc-jena.mpg.de (N.C.); mmahecha@bgc-jena.mpg.de (M.D.M.); mreichstein@bgc-jena.mpg.de (M.R.)

<sup>2</sup> Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

<sup>3</sup> Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, Droevendaalsesteeg 3, 6708 PB Wageningen, The Netherlands; E-Mail: jan.verbesselt@wur.nl

<sup>4</sup> Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; E-Mail: christopher.s.neigh@nasa.gov

\* Author to whom correspondence should be addressed; E-Mail: mforkel@bgc-jena.mpg.de; Tel.: +49-3641-576-283; Fax: +49-3641-577-200.

*Received: 27 February 2013; in revised form: 17 April 2013 / Accepted: 25 April 2013 /*

*Published: 3 May 2013*



## 2 Trend change detection in NDVI time series: effects of inter-annual variability and methodology

### Abstract

Changing trends in ecosystem productivity can be quantified using satellite observations of Normalized Difference Vegetation Index (NDVI). However, the estimation of trends from NDVI time series differs substantially depending on analyzed satellite dataset, the corresponding spatiotemporal resolution, and the applied statistical method. Here we compare the performance of a wide range of trend estimation methods and demonstrate that performance decreases with increasing inter-annual variability in the NDVI time series. Trend slope estimates based on annual aggregated time series or based on a seasonal-trend model show better performances than methods that remove the seasonal cycle of the time series. A breakpoint detection analysis reveals that an overestimation of breakpoints in NDVI trends can result in wrong or even opposite trend estimates. Based on our results, we give practical recommendations for the application of trend methods on long-term NDVI time series. Particularly, we apply and compare different methods on NDVI time series in Alaska, where both greening and browning trends have been previously observed. Here, the multi-method uncertainty of NDVI trends is quantified through the application of the different trend estimation methods. Our results indicate that greening NDVI trends in Alaska are more spatially and temporally prevalent than browning trends. We also show that detected breakpoints in NDVI trends tend to coincide with large fires. Overall, our analyses demonstrate that seasonal trend methods need to be improved against inter-annual variability to quantify changing trends in ecosystem productivity with higher accuracy.

## 2.1 Introduction

Climate change will likely change biome distributions, ecosystem productivity and forest carbon stocks (Lucht et al., 2006). Such ecosystem changes can be detected and quantified using multi-temporal satellite observations of the land surface. Different states of the land surface can be measured by satellite-derived biophysical parameters (Coppin et al., 2004). The Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) is a remotely-sensed measure of vegetation greenness and is related to structural properties of plants – like leaf area index (Turner et al., 1999) and green biomass (Gamon et al., 1995) – but also to properties of vegetation productivity – like absorbed photosynthetic active radiation and foliar nitrogen (Fensholt et al., 2004; Gamon et al., 1995). As NDVI is related to such a variety of vegetation properties, multiple explanations for a change in NDVI signals are possible. Nevertheless, the NDVI from AVHRR (Advanced Very High Resolution Radiometer) satellite observations is the only global vegetation dataset which spans a time period of three decades and thus allows the quantification and attribution of ecosystem changes as a result of ecosystem dynamics and varying climate conditions. Different ecosystem changes can be analyzed from NDVI time series. For example, annual mean or peak NDVI provides an integrated view on photosynthetic activity (Myneni et al., 1995), the seasonal NDVI amplitude is related to the composition of evergreen and deciduous vegetation (DeFries et al., 1995) and the length of the NDVI growing season can be related to phenological changes (Tucker et al., 2001). Thus, trend detection in NDVI time series can help to identify and quantify recent changes in ecosystem properties from a local to global scale.

Indeed, positive NDVI trends (“greening”) occur in the high latitudes (Myneni et al., 1997a). These greening trends were reproduced by a Dynamic Global Vegetation Model (DGVM) and attributed to increasing temperatures (Lucht et al., 2002). The temperature increase drives an expansion of shrubs in the arctic Tundra, which can be observed as greening trends (Myers-Smith et al., 2011). The initial greening trend stalled or reversed in large parts of the boreal forest of Northern America. Negative NDVI trends (“browning”) are associated with fire activity (Goetz et al., 2005), increasing water vapour pressure deficit (Bunn et al., 2007) or to cooling spring temperatures (Wang et al., 2011). Regional changes in summer precipitation changed

greening NDVI trends to browning trends also in Eurasian boreal forests in the late 1990s (Piao et al., 2011). Nevertheless, browning NDVI trends are highly discussed because they differ based on the used satellite dataset (Beck and Goetz, 2011; Fensholt and Proud, 2012). Most studies used the GIMMS (Global Inventory, Monitoring, and Modeling Studies) NDVI dataset which was produced based on 4 km AVHRR satellite observations (Tucker et al., 2005). A comparison between the GIMMS dataset and a Canadian dataset shows weaker post-fire recovery trends and more negative NDVI trends in unburned forests in the GIMMS dataset (Alcaraz-Segura et al., 2010). Other studies confirm trend estimates based on the GIMMS dataset: Despite of some regional differences in areas at very high latitudes with low vegetation cover, NDVI trends from the GIMMS dataset agree with trends from MODIS data (Moderate Resolution Imaging Spectrometer) (Beck and Goetz, 2011; Fensholt and Proud, 2012; Parent and Verbyla, 2010). Trends from the GIMMS dataset compare well with trends computed from Landsat imagery (Fraser et al., 2011). Changes in tree rings (Beck et al., 2011; Berner et al., 2011), temperature-induced drought stress or insect disturbances (Verbyla, 2011) were also observed in regions with browning NDVI trends. In fact, impacts of recent trends and variability of climate on ecosystems can be observed using long-term NDVI time series.

The estimation of trends depends on the length, temporal and spatial resolution of the time series, the quality of the measured data (Sulkava et al., 2007) and the used statistical method. Many studies calculated trends based on annual time steps from annually or seasonally aggregated values using regression analysis (Eklundh and Olsson, 2003). However, the use of linear regression analysis for estimating trends in NDVI time series violates statistical assumptions such as the independence of observations, due to temporal autocorrelation or homogeneity (de Beurs and Henebry, 2004b). Accordingly, the application of temporal autocorrelation structures (Goetz et al., 2005) or the use of the non-parametric Mann-Kendall test on NDVI time series was suggested to circumvent the limitations of regression analysis (de Beurs and Henebry, 2004b, 2010a). The annual aggregation of time series for trend analysis reduces the temporal resolution and time series length. The time series length is critical in determining the significance of the trend in a statistical test. On the other hand, annual aggregation supports the analysis of trends by eliminating the seasonal cycle in the

NDVI time series. The seasonal cycle introduces a seasonal correlation structure that hampers trend analysis. In this context, methods were developed that make use of the full resolution time series by estimating and subtracting the seasonal cycle or by modelling the seasonal signal (de Jong et al., 2011a; Mahecha et al., 2010a; Verbesselt et al., 2010a, 2010b; Wu et al., 2008). Overall, trend estimates from these different methods result in similar general spatial patterns of the major regional greening and browning trends but substantial differences in areas with weak trends (de Jong et al., 2011a). In short, all trend estimation methods embed caveats that may be more or less critical depending on the application.

NDVI trends are not always monotonic but can change. A positive trend can change for example into a negative one and vice versa. Changes between initial greening trends in the 1980s to browning trends from the 1990s onwards in high latitude regions were detected based on the GIMMS NDVI dataset (de Jong et al., 2011b). NDVI trend changes of this kind can be either gradual or abrupt (de Jong et al., 2013b). For example, increasing temperatures in temperature-limited ecosystems can first support vegetation growth that results in greening NDVI trends, while a further warming can induce drought stress that slightly turns the initial greening to a browning trend (gradual change). Disturbances such as fire events can reduce the NDVI signal and initiate post-fire recovery that results in a greening trend (abrupt change). Recently, statistical methods were developed and applied to NDVI time series to detect such changes (called breakpoints) in trends. Methods like BFAST (Breaks for additive seasonal and trend) (Verbesselt et al., 2010a, 2010b) combine trend estimation with approaches that account for breakpoints in the trend. However, the reliability of such breakpoints in NDVI trends in high latitude regions is not yet assessed.

Breakpoints in NDVI time series are related to different effects caused by inter-annual variability. Inter-annual variability of NDVI time series can be caused by (1) artefacts of a harmonized dataset from different sensors, (2) meteorological distortions like clouds or snow cover and (3) environmental processes like effects of year-to-year variations in weather conditions on plant activity or ecosystem disturbances. Inter-annual variability affects the annual mean (e.g., reduction of NDVI because of a disturbance), seasonality (e.g., longer growing season because of prolonged warmer temperatures) as well as short-term patterns (e.g., unusual snowfall in a summer month) of NDVI

time series. The aggregation of NDVI time series to mean annual values integrates these different effects which, despite the loss in temporal detail, allow us to define and quantify inter-annual variability as the standard-deviation of mean annual NDVI values.

The purpose of this study is to evaluate the performance of different methods for detecting trends and trend breakpoints in long-term NDVI time series. Previous studies have used different trend and breakpoint analysis methods on NDVI time series without or with limited demonstration of its methodological robustness (Wang et al., 2011). By evaluating the performance of such methods, this study will enable a critical appraisal of combined trend and breakpoint detection methods for their application on NDVI time series. The methods chosen for evaluation differ on their used temporal resolution of NDVI time series, how seasonality is accounted for, and how the trend is estimated. All approaches make use of the same breakpoint detection algorithm (Bai and Perron, 2003). A factorial experiment was performed based on surrogate (or “artificial”) NDVI time series with different levels of trend magnitude, inter-annual and short-term variability, seasonal amplitudes and a varying number of trend changes. We tested whether the methods are able to re-detect the prescribed trend (i.e., slope of the trend) and trend changes (i.e., number and timing of breakpoints) in the surrogate time series. Additionally, methods were applied to real NDVI time series of Alaska and the plausibility of breakpoints was assessed in comparison to fire events and drought periods. Our results reveal a clear dependence of the method’s performance on the degree of inter-annual variability.

## **2.2 Data and methods**

### **2.2.1 GIMMS NDVI3g dataset**

The GIMMS NDVI3g dataset (third generation GIMMS NDVI) is a newly available long-term NDVI dataset and was derived from NOAA AVHRR data (National Oceanic and Atmospheric Administration, Advanced Very High Resolution Radiometer) (Pinzon and Tucker, 2014; Xu et al., 2013). In comparison to a previous version of the dataset (Tucker et al., 2005), it was improved for applications in high-latitude regions through calibrations to stable targets in these regions (Pinzon and Tucker, 2014; Xu et

al., 2013). The dataset covers the period July 1981 until December 2011 with a 2-weekly temporal and 8 km spatial resolution. The quality of AVHRR data is affected from sensor changes between the NOAA satellites and orbital decay but it was shown that trends based on the previous NDVI dataset are not affected by these artefacts (Kaufmann et al., 2000). In the new GIMMS NDVI3g dataset such effects were substantially reduced (Pinzon and Tucker, 2014; Xu et al., 2013).

We further pre-processed the GIMMS NDVI3g dataset for the specific requirements of our study. The year 1981 was excluded from our analysis in order to analyse only years with full data coverage. Especially in high-latitude regions NDVI observations are often affected from snow or cloud cover. Such NDVI values are flagged as “snow” or “interpolated” in the GIMMS NDVI3g dataset (Pinzon and Tucker, 2014; Xu et al., 2013). The reliability of such interpolated NDVI values under snow or cloud conditions is unclear. We addressed this fact in two ways: (1) we assumed interpolated NDVI observations under snow conditions are the best available estimate and we did not change these NDVI values (hereinafter called “all” observations). (2) NDVI values that were flagged as “snow” were excluded from the analysis (hereinafter called “ex” observations). We kept interpolated observations that were not additionally flagged as “snow” to make sure to use enough observations throughout the growing season because three of the four assessed trend methods need seasonal observations in order to be applicable for trend detection (Section 2.2.3). To exclude potentially remaining effects of cloud or haze contaminations, the dataset was further aggregated to monthly temporal resolution using the monthly maximum value which is a commonly applied procedure (Holben, 1986).

### **2.2.2 Breakpoint detection algorithm**

The breakpoint detection algorithm as described by Bai and Perron (2003) and Zeileis *et al.* (2003) was used in this study. The breakpoint detection algorithm searches for a structural change in a regression relationship, *i.e.*, for varying regression parameters before and after the breakpoint. That means a detected breakpoint splits a time series in two segments. In a first step, the ordinary-least squares moving sum (MOSUM) test is used to test for the existence of breakpoints in the time series. If the test indicates significant structural changes ( $p \leq 0.05$ ), different numbers and locations of breakpoints

are iteratively tested in the second step. For this purpose, the optimal number of breakpoints is estimated by minimizing the Bayesian Information Criterion (BIC). The optimal position of a breakpoint is estimated by minimizing the residual sum of squares of this regression (Bai and Perron, 2003; Zeileis et al., 2003).

The breakpoint detection algorithm was used based on recommendations of Bai and Perron (2003). In order to detect long-term trend changes, a minimum amount of observations between two breakpoints was defined as 48 monthly observations (or four years) and a maximum number of two allowed breakpoints were selected. Therefore, an optimized number of breakpoints between zero and two can be detected. This prevented that detected breakpoints are solely affected by year-to-year changes and supported the detection of only major breakpoints in the long-term trend. Further, detected time series segments of a length smaller than eight years were not considered as trends.

### **2.2.3 Methods for trend estimation**

#### *2.2.3.1 Trend estimation on annual aggregated time series (Method AAT)*

Method AAT estimates trends and trend changes on annual aggregated time series. The seasonal NDVI time series is first aggregated to annual values. The annual mean, growing season mean or annual peak NDVI can be calculated to aggregate the seasonal NDVI time series to annual values. Mean annual NDVI was used for the factorial experiment based on surrogate time series. Breakpoints are estimated on the annual time series using the method of Bai and Perron (2003). For each derived trend segment the slope of the trend is estimated by linear least-squares regression of the annual values against time. The significance of the trend in each time series segment is estimated by the Mann-Kendall trend test applied on the annual aggregated NDVI values (Mann, 1945).

#### *2.2.3.2 Trend estimation based on a season-trend model (Method STM)*

The trend and breakpoint estimation in method STM (season-trend model) is based on the classical additive decomposition model and we followed the formulation used in BFAST (Verbesselt et al., 2010b, 2012). The full temporal-resolution NDVI time series is

explained by a piecewise linear trend and a seasonal model in a regression relationship. Thus, the NDVI value  $y$  at a time  $t$  can be expressed as:

$$y_t = \alpha_1 + \alpha_2 t + \sum_{j=1}^k \gamma_j \sin\left(\frac{2\pi j t}{f} + \delta_j\right) + \varepsilon_t \quad (2.1)$$

where  $\alpha_1$  is the intercept and  $\alpha_2$  the slope of the trend,  $\gamma$  are the amplitudes and  $\delta$  the phases of  $k$  harmonic terms and  $\varepsilon$  is the residual error (Verbesselt et al., 2012). The frequency  $f$  is the number of observation per year (*i.e.*, 12 for monthly observations). Parameters  $\alpha_1$ ,  $\alpha_2$  are estimated using ordinary least squares (OLS) regression whereby the derived time series segments are considered as categorical interaction term with the trend slope  $\alpha_2$ . The significance of the trend in each segment is estimated from a t-test on the interaction parameter of the regression between time series segment and  $\alpha_2$ .

### 2.2.3.3 Trend estimation on de-seasonalized time series

Methods MAC (mean annual cycle) and SSA (annual cycle based on singular spectrum analysis) estimate trends on seasonal-adjusted time series, which is the full-resolution time series with removed seasonality. The seasonal-adjusted time series  $a$  is the difference between the original NDVI time series  $y$  and the seasonal cycle  $s$ :

$$a = y - s \quad (2.2)$$

The slope of the trend  $\alpha_2$  is estimated using OLS from the seasonal-adjusted time series:

$$a_t = \alpha_1 + \alpha_2 t + \varepsilon_t \quad (2.3)$$

Breakpoints are estimated on Equation (3) with different regression coefficients for each trend segment. The significance of the trend in each time series segment is estimated with the Mann-Kendall trend test applied on the seasonal-adjusted time series.

The seasonal cycle (or annual cycle)  $s$  (Equation (2)) is represented by a mean annual cycle (method MAC) and by a modulated annual cycle (method SSA). The MAC is estimated as the mean seasonal cycle from the seasonal cycles of all years. This implies that each year has the same amplitude and frequency in the seasonal cycle. However, the concept of a fixed seasonal cycle is questionable as it can change due to external forcing (Wu et al., 2008). For example, phenological cycles might change without affecting the overall trend in a time series. Therefore, method SSA is based on a modulated annual cycle with slightly varying frequencies and amplitudes of the



seasonal cycle amongst years. The modulated seasonal cycle was estimated using a one-dimensional singular spectrum analysis (SSA) as described in (Golyandina et al., 2001). Singular spectrum analysis was previously used to separate remotely-sensed FAPAR (fraction of observed photosynthetically active radiation) time series into low and high frequency and seasonal time series components (Mahecha et al., 2010a). SSA decomposes in a first step a time series into different sub-signals with characteristic frequencies. In a second step, the sub-signals with an annual frequency were summed to build up the modulated annual cycle.

## **2.2.4 Simulation of surrogate time series**

### *2.2.4.1 Estimation of inter-annual variability, seasonality and short-term variability from observed time series*

An important aspect of the experimental design was the prescription of time series properties in the surrogate (artificial) data that were observed in the NDVI datasets. In order to create surrogate time series that mimic the full range of possible real world data, the mean, trend, inter-annual variability, seasonality and short-term variability was estimated for all observed NDVI time series of Alaska in a simple step-wise approach (Figure 2.1):

- (1) The mean of each NDVI time series was calculated.
- (2) In the second step, monthly values were averaged to annual values and the trend was calculated according to method AAT but without computing breakpoints. Hence, the slope of the annual NDVI trend over the full length of the time series was estimated.
- (3) To estimate the inter-annual variability, the standard deviation and range of the annual anomalies were calculated. The mean of the time series and the derived trend component from step (2), were subtracted from the annual values to derive the trend-removed and mean-centred annual values (annual anomalies). If the trend slope was not significant ( $p > 0.05$ ), only the mean was subtracted. The standard deviation and the range of the annual anomalies were computed as measures for the inter-annual variability of the time series.
- (4) In the next step, the range of the seasonal cycle was estimated. The mean, the trend component and the annual anomalies were subtracted from the original time series to

calculate a detrended, centered and for annual anomalies adjusted time series. Based on this time series the seasonal cycle was estimated as the mean seasonal cycle and the range was computed.

(5) In the last step, the standard deviation and the range of the short-term anomalies were computed. Short-term anomalies were computed by subtracting the mean, the trend component, the annual anomalies and the mean seasonal cycle from the original time series. The result is the remainder time series component. The standard deviation of the remainder time series component is a measure of short-term variability.

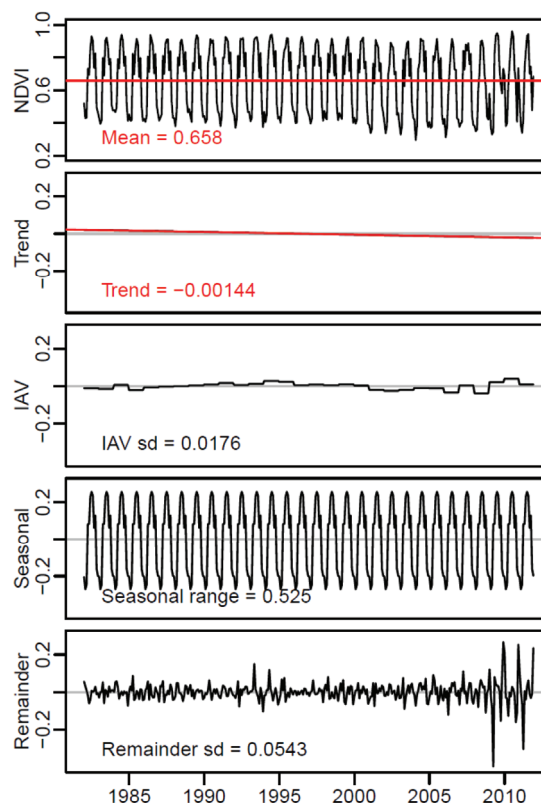


Figure 2.1: Estimated time series components for a random-selected example grid cell in central Alaska (3\*3 grid cells averaged around central pixel 146.424°W, 64.762°N). The upper panel shows the original Normalized Difference Vegetation Index (NDVI) time series with its mean value (red line). The next panels show the estimated trend, inter-annual variability (IAV) (i.e., annual anomalies), seasonality (i.e., mean seasonal cycle) and short-term variability (remainder component), respectively. The sum of mean, trend, IAV, seasonal and remainder component equals the original time series.

All the described time series properties were estimated on the full NDVI dataset including all observations (i.e., including snow-affected observations). Hence, we could generate a wide range of gap free surrogate time series. This procedure was applied for all NDVI time series of Alaska to estimate spatial and statistical distributions of the

mean NDVI, the overall trend slope, the inter-annual variability as the standard deviation of the annual anomalies, the range of the seasonal cycle and the short-term variability as the standard deviation of the remainder time series (Figure 2.2).

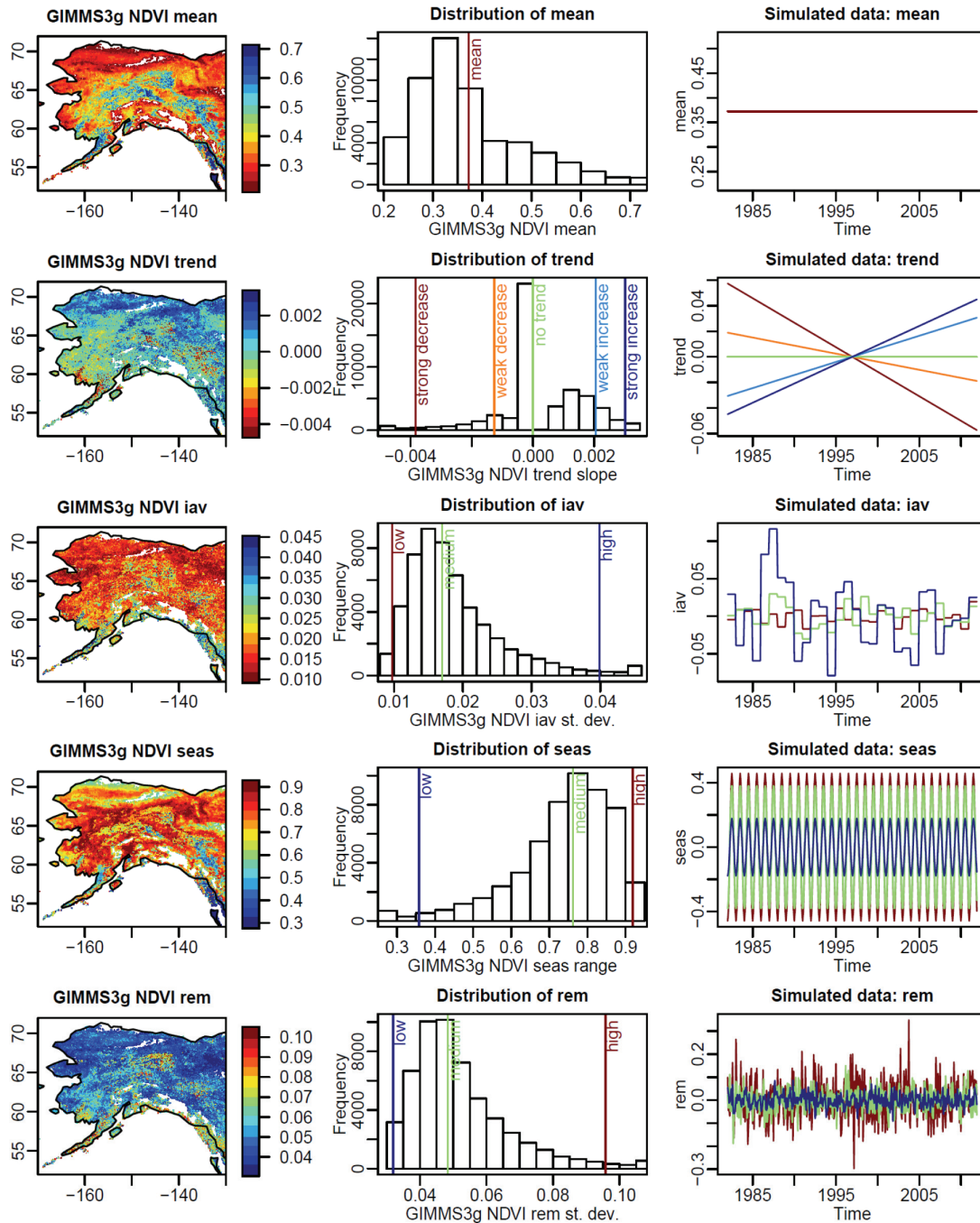


Figure 2.2: Spatial and statistical distributions of NDVI time series properties in Alaska and time series components of the simulated NDVI time series.

The left panel shows from top to bottom maps of the following time series properties: mean annual NDVI, slope of the annual trend ( $\Delta\text{NDVI}/\text{year}$ ), standard deviation of the inter-annual variability (iav), range of the seasonal cycle (seas), and the standard deviation of the remainder component (rem). The middle panel shows the statistical distribution of these properties, respectively. The right panel shows examples of the surrogate time series components.

#### 2.2.4.2 *Surrogate time series and factorial experiment*

Surrogate time series were simulated based on addition of different time series components that were estimated from observed time series properties:

$$y_t = m + T_t + I_t + S_t + R_t \quad (2.4)$$

where  $m$  is the mean,  $T$  is the trend component value,  $I$  the inter-annual variability component value,  $S$  the seasonal component and  $R$  the remainder component at time step  $t$ . As the estimated values of  $T$  and  $I$  have an annual temporal resolution, they were repeated for each time step  $t$  of the same year (forming a step function) to create the simulated monthly time series. The mean was taken from the mean of the observed distribution of average NDVI. We selected only one mean value for all surrogate time series because differences in mean are expressed by the intercept of the linear regression models and will not affect the trend estimate. For each of the other components, different levels were used to create surrogate time series:

- (1) Trend: Time series with strong and weak positive, strong and weak negative and without a trend were created. Different magnitudes of trend slopes were derived from the 1% percentile of the observed distribution of trend slopes (strong decrease), 25% percentile (weak decrease), median (no trend), 75% percentile (weak increase) and 99% percentile (strong increase), respectively.
- (2) Inter-annual variability: Time series with low, medium and high inter-annual variability were created based on normal-distributed random values with zero mean and a standard deviation according to the 1%, 50% and 99% percentiles of the observed distribution of the standard deviation of annual anomalies. Values outside the observed ranges of inter-annual variability were set to the minimum or maximum of the observed distribution, respectively.
- (3) Seasonality: Seasonal cycles based on a harmonic model with low, medium, and high amplitudes were created according to the observed 1%, 50% and 99% percentiles of the distribution of seasonal ranges.
- (4) Short-term variability: Different levels of short-term variability were created based on normal-distributed random values with zero mean and a standard deviation according to the 1%, 50% and 99% percentiles of the observed distribution of the standard deviation of remainder time series values.

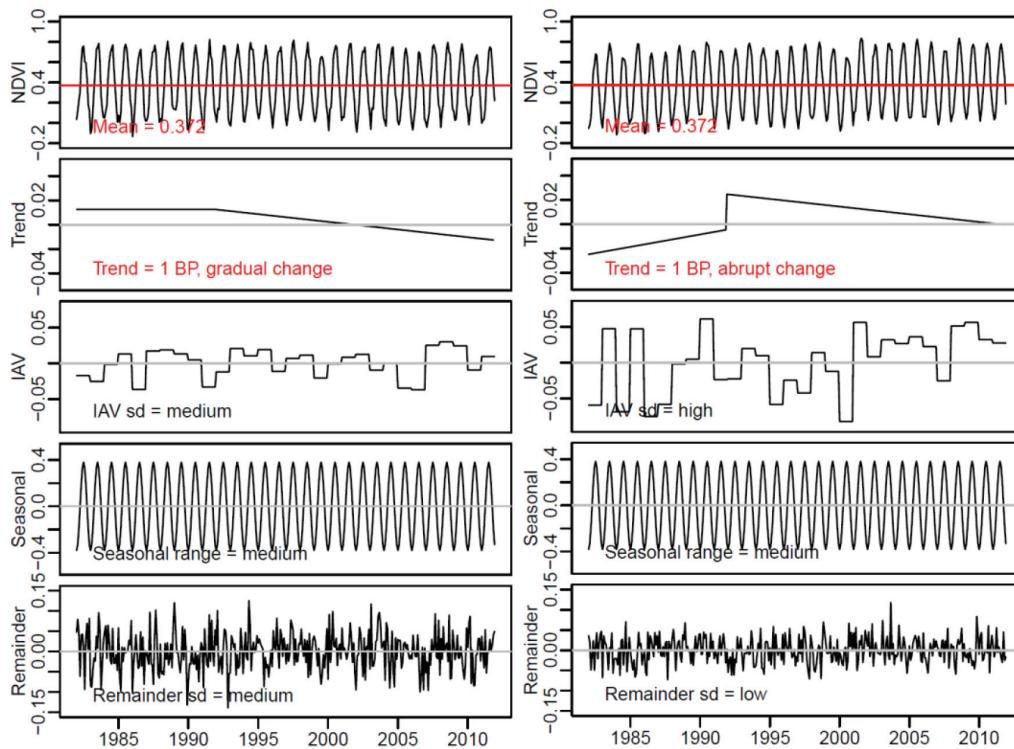


Figure 2.3: Examples of simulated time series with different components of trend, IAV, seasonal and remainder referring to the simulated trend, inter-annual variability, seasonal and remainder time series components, respectively.

The sum of these time series components gives the total simulated surrogate NDVI time series (upper panel). Left: time series with one breakpoint and gradual change (e.g., caused by gradual changes in environmental conditions), no trend in the first segment and decreasing trend in the second segment, medium inter-annual variability, medium seasonality and medium short-term variability. Right: Time series with one breakpoint and abrupt change (e.g., caused by a few years with exceptional favourable growing conditions), increasing trend in first segment and decreasing trend in second segment, high inter-annual variability, medium seasonality and low short-term variability.

To introduce trend changes in these surrogate time series, trend components with one and two breakpoints as well as gradual or abrupt changes were created. In case of one breakpoint, the break was introduced 120 months after the beginning of the time series and in case of two breakpoints after 107 and 215 months, respectively. That means one time series can have one to three time series segments with a length of 360 months (30 years) in case of no breakpoint, 120 and 241 months in case of one breakpoint, respectively, and 107, 107 and 146 months in case of two breakpoints, respectively. The type of trend change was considered as an additional factor, whereby gradual change and abrupt changes were distinguished. A gradual change is a change between two trend segments, for which the last value of the trend component in the first segment equals the first value of the following trend segment (Figure 2.3). In case of abrupt

changes, trend components are disconnected between two segments (Figure 2.3). In summary, the following factors with different levels were considered in the experiment:

- (1) Type of trend and number of breakpoints/segments (maximum 2 breakpoints = maximum 3 segments per time series with positive, negative or no trend = 27 possibilities),
- (2) Trend magnitude (weak, strong),
- (3) Inter-annual variability (low, medium, high),
- (4) Short-term variability (low, medium, high),
- (5) Type of trend change (gradual, abrupt) and
- (6) Range of seasonal cycle (low, medium, high).

For each combination of these factors one surrogate time series was created. Because some combination are physically not possible (e.g., abrupt or gradual change but 0 breakpoints), in total 1377 surrogate time series were created.

## **2.2.5 Evaluation of method performances**

### *2.2.5.1 Evaluation of breakpoints*

To evaluate the performance of the methods regarding the estimated breakpoints, the difference in number and timing of breakpoints were compared. The difference between the estimated number of breakpoints and the number of real breakpoints was calculated. The number of real breakpoints is the amount of breakpoints that was used to simulate the surrogate time series.

The timing of an estimated breakpoint was compared against the timing of the real breakpoint. For each estimated breakpoint the nearest real breakpoint was selected and the absolute temporal difference (in months) between them was calculated. If the difference is larger than five years the real breakpoint was set as undetected.

### *2.2.5.2 Evaluation of trend slopes and significances*

In order to evaluate the direction and significance of an estimated trend, estimated trends were compared with the real trend in a trend segment of the simulated time series. The slope and p-value of the real trend was calculated for each method based on the known real breakpoints and time series segments. To compare direction and

significance of trends, trend slopes and p-values of real and estimated trends in a time series segment were classified in six trend classes:

N3: significant negative trend (slope < 0 and  $p \leq 0.05$ )

N2: non-significant negative trend (slope < 0 and  $0.05 < p \leq 0.1$ )

N1: no trend with negative tendency (slope < 0 and  $p > 0.1$ )

P1: no trend with positive tendency (slope > 0 and  $p > 0.1$ )

P2: non-significant positive trend (slope > 0 and  $0.05 < p \leq 0.1$ )

P3: significant positive trend (slope > 0 and  $p \leq 0.05$ ).

Confusion matrices of estimated and prior trend classes were computed for each method in order to evaluate the accuracy of the methods for trend estimation. Confusion matrices (alternatively called contingency table or error matrix) are standard tools to compare errors between two classifications (Foody, 2002). Congalton et al. (1991) suggested to normalize confusion matrices in order to eliminate effects of different sample sizes per class and to make confusion matrices between different classifications, i.e., different trend methods, comparable. Iterative Proportional Fitting Procedure (IPFP) (Deming and Stephan, 1940) was used to normalize confusion matrices to row and column (marginal) totals of 100%. Based on normalized confusion matrices, the total normalized accuracy and the Kappa coefficient were calculated (Congalton, 1991; Foody, 2002) to quantify the performance of methods for trend estimation. The total accuracy ranges between 0% (worst accuracy) and 100% (complete agreement of the two classifications) and the Kappa coefficient ranges between 0 (worst agreement) and 1 (complete agreement).

### 2.2.5.3 *Evaluation of the overall performance for trend and breakpoint estimation*

The overall performance of a method was quantified by comparing the estimated with the real trend component (Equation 2.4). For this purpose, the root mean square error (RMSE) between the estimated and the real trend component  $T$  was calculated for the total length  $n$  of the simulated time series:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (T_{est} - T_{real})^2}{n}} \quad (2.5)$$

This formulation involves both an effect of the estimated trend and the estimated breakpoints. An analysis-of-variance (ANOVA) was calculated for the RMSE in order to identify the factors that explain the highest fraction of the RMSE. Trend magnitude, inter-annual variability, seasonality, short-term variability, type of trend change, number of real breakpoints and method as well as their second-order interactions were used as explanatory variables in the ANOVA.

## 2.2.6 Application to real time series of Alaska: ensemble of NDVI trends

All trend and breakpoint methods were applied to real NDVI time series of Alaska and parts of Yukon to assess differences between methods based on real data. The application of the four methods (AAT, STM, MAC, SSA) allows creating a multi-method ensemble of NDVI trend estimates. As many NDVI observations in northern regions are affected from snow, clouds or other distortions, the use of such poor quality observations causes additional uncertainties in NDVI trend estimates. To account for the effect of snow-affected observation, all methods were applied on the NDVI time series with all observations (“all”) and on the NDVI time series excluding snow-affected values (“ex”) (see Section 2.2.1). Additionally, method AAT was applied on the annual peak NDVI (defined as the annual quantile 0.9) to analyse trends (called AAT-peak) because vegetation growth in high-latitude ecosystems is usually limited to the peak production period. This setup of trend methods on the real dataset resulted in nine trend and breakpoint estimates for Alaska (AAT-all, AAT-ex, AAT-peak, STM-all, STM-ex, MAC-all, MAC-ex, SSA-all, SSA-ex). From all the nine trend estimates, the ensemble mean and standard deviation of the number of detected breakpoints, of the duration of greening and browning trends and of the trend slope were calculated. The ensemble mean NDVI slope  $\alpha$  was calculated from the weighted mean slope  $\alpha_m$  of a method  $m$ , weighted by the length  $l$  of the corresponding time series segment  $seg$  and the p-value  $p$  of the trend in the segment expressed as significance:

$$\alpha = \frac{1}{n_{method}} \sum_{m=1}^{m=9} \alpha_m \quad (2.6)$$



$$\alpha_m = \frac{\sum_{seg=1}^{seg=nseg} \alpha_{m,seg} \times (l_{seg} \times (1 - p_{seg}))}{\sum_{seg=1}^{seg=nseg} (l_{seg} \times (1 - p_{seg}))} \quad (2.7)$$

Time series segments with a length  $l$  shorter than eight years were excluded from this analysis and remaining segment lengths (between eight and 30 years) were scaled to 0-1 before using them as weights in Equation 2.7. The uncertainty of NDVI trend slopes was calculated as the standard deviation of mean slopes  $\alpha_m$ .

Additional datasets were used to assess the plausibility of detected breakpoints. Fire perimeter observations from the Alaskan Large Fire Database (Frames, 2012; Kasischke et al., 2002) were compared against the spatial distribution and timing of detected breakpoints. Gridded precipitation time series from the GPCC dataset (Global Precipitation Climatology Center) (Schneider et al., 2008) and temperature time series from the CRU dataset (Mitchell and Jones, 2005) were used to compare breakpoints with annual temperature and precipitation anomalies (baseline 1982–2009). Additionally, photos taken in 2008 at burnt areas of the year 2004 were compared with detected breakpoints in NDVI time series to visually inspect the post-fire vegetation status in NDVI time series with breakpoints.

## 2.3 Results

### 2.3.1 Observed and simulated properties of NDVI time series

To simulate surrogate time series, statistical distributions of NDVI mean, trend, seasonality, inter-annual and short-term variability were computed from observed NDVI time series in Alaska (Figure 2.2). Mean NDVI ranged between 0.2 and 0.87 with an average value of ca. 0.37. The highest NDVI means occurred in the central Alaskan boreal forest and the lowest values in the northern Tundra regions (Figure 2.2). The mean of 0.37 was used as the mean value in all simulated time series. NDVI trend slopes ranged from  $-0.0047$  to  $0.0034 \Delta\text{NDVI}/\text{year}$ , with the lowest values in some boreal forest regions and the highest values in the northern Tundra regions. NDVI trend slopes of  $-0.0038$  (strong decrease),  $-0.0013$  (weak decrease),  $0$  (no trend),  $0.002$  (weak increase) and  $0.003 \Delta\text{NDVI}/\text{year}$  (strong increase) were used to create trend

components for the surrogate time series. The standard deviation of the annual averaged NDVI values was used as a measure for inter-annual variability. It ranged between 0.009 in some regions and 0.045 in south-western Alaska. The 0.01, 0.5, and 0.99 quantiles of 0.0097, 0.016 and 0.039 were used to create surrogate NDVI time series with low, medium and high inter-annual variability, respectively (Figure 2.2). The range of the mean seasonal cycle ranged from 0.27 in southern coastal regions of Alaska to 0.93 in some northern Tundra regions. Seasonal ranges of 0.34 (low), 0.76 (medium) and 0.91 (high) were used to create surrogate time series. The standard deviation of the remainder time series component was used a measure of short-term variability and ranges between 0.03 and 0.1. Values of 0.031 (low), 0.048 (medium) and 0.096 (high) were used as the standard-deviation of normal distributed random number to create surrogate time series components of short-term variability (Figure 2.2). Because of the fact, that the percentiles 1% and 99% were chosen as the low and high levels for inter-annual variability, seasonality and short-term variability, simulated time series can be strongly dominated by seasonality (e.g., in case of high seasonal range but low inter-annual and short-term variability) or can show almost random behaviour (e.g., in case of low seasonality but high inter-annual and short-term variability). Consequently, the simulated NDVI time series covered not only a wide and extreme range of observed time series properties of the study region but contain time series properties that might occur under different environmental conditions.

### **2.3.2 Evaluation of estimated breakpoints**

To evaluate the performance of the methods to detect breakpoints in trends, the real and estimated breakpoints were compared. For this purpose, the difference between the estimated and real number of breakpoints was calculated and analysed grouped by methods and factors (Figure 2.4). All methods underestimated the number of breakpoints, i.e., the number of false positive detected breakpoints was small (0.36% for method AAT, 3.8% for STM, 15.5% for MAC and 15.3% for SSA). Method AAT did not detect any breakpoint if there was no breakpoint whereas methods STM, MAC and SSA detected in up to 35% of all cases one or two breakpoints if there was no breakpoint in the surrogate time series (Figure 2.4 e). The range of the seasonal cycle

had no effect on the estimation of the number of breakpoints (Figure 2.4 g). The performance of all methods to estimate the correct number of breakpoints depended also on inter-annual and short-term variability. The performance of method AAT did not change with increasing inter-annual and short-term variability while methods STM, MAC and SSA had an increasing number of false positive detected breakpoints with increasing inter-annual variability and a slightly decreasing number of false positive detected breakpoints with increasing short-term variability (Figure 2.4 c-d). All methods tended to perform better in case of abrupt trend changes than in case of gradual trend changes (Figure 2.4 f). In short, the difference between estimated and real number of breakpoints depended mostly on the number of real breakpoints as well as short-term and inter-annual variability.

Furthermore, the temporal difference between a detected and the closest real breakpoint was calculated, to evaluate the performance of methods regarding the timing of breakpoints (Figure 2.5). The difference in timing was not calculated if a method did not detect a breakpoint although there were real breakpoints. In average, method AAT performed better (mean absolute difference in breakdates = 9.2 months) than method STM (16 months) and methods MAC and SSA (both 19 months) (Figure 2.5 a). The error of the timing of breakpoints increased with increasing inter-annual variability (Figure 2.5 c). Increasing short-term variability resulted only for method AAT in an increasing timing error (Figure 2.5 d). All methods had a better timing of breakpoints in case of abrupt changes (Figure 2.5 f). The difference in breakdates was lower in case of multiple breakpoints (Figure 2.5 e). The range of the seasonal cycle has no effects on the timing of breakpoints (Figure 2.5 f). In summary, the correct timing of breakpoints depended mostly on inter-annual variability and the type of trend change.

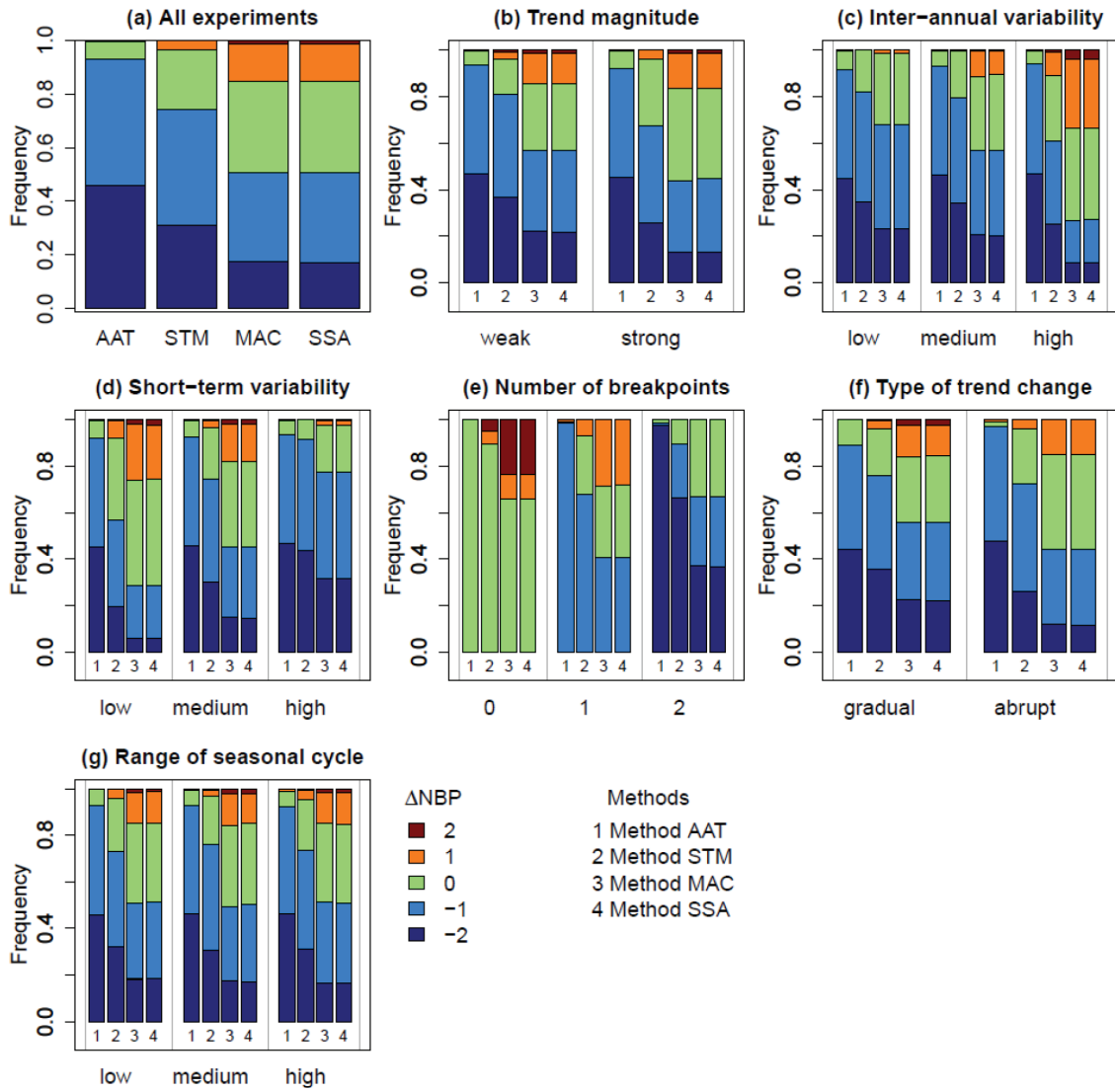


Figure 2.4: Frequencies of differences between real and estimated number of breakpoints for the different methods from all experiments.

Blue indicates underestimation, red overestimation of the number of real breakpoints. (a) Performance of the methods in all experiments. (b) Grouped by trend magnitude. (c) Grouped by inter-annual variability. (d) Grouped by short-term variability. (e) Grouped by the real number of breakpoints. (f) Grouped by the type of trend change. (g) Grouped by the range of the seasonal cycle.

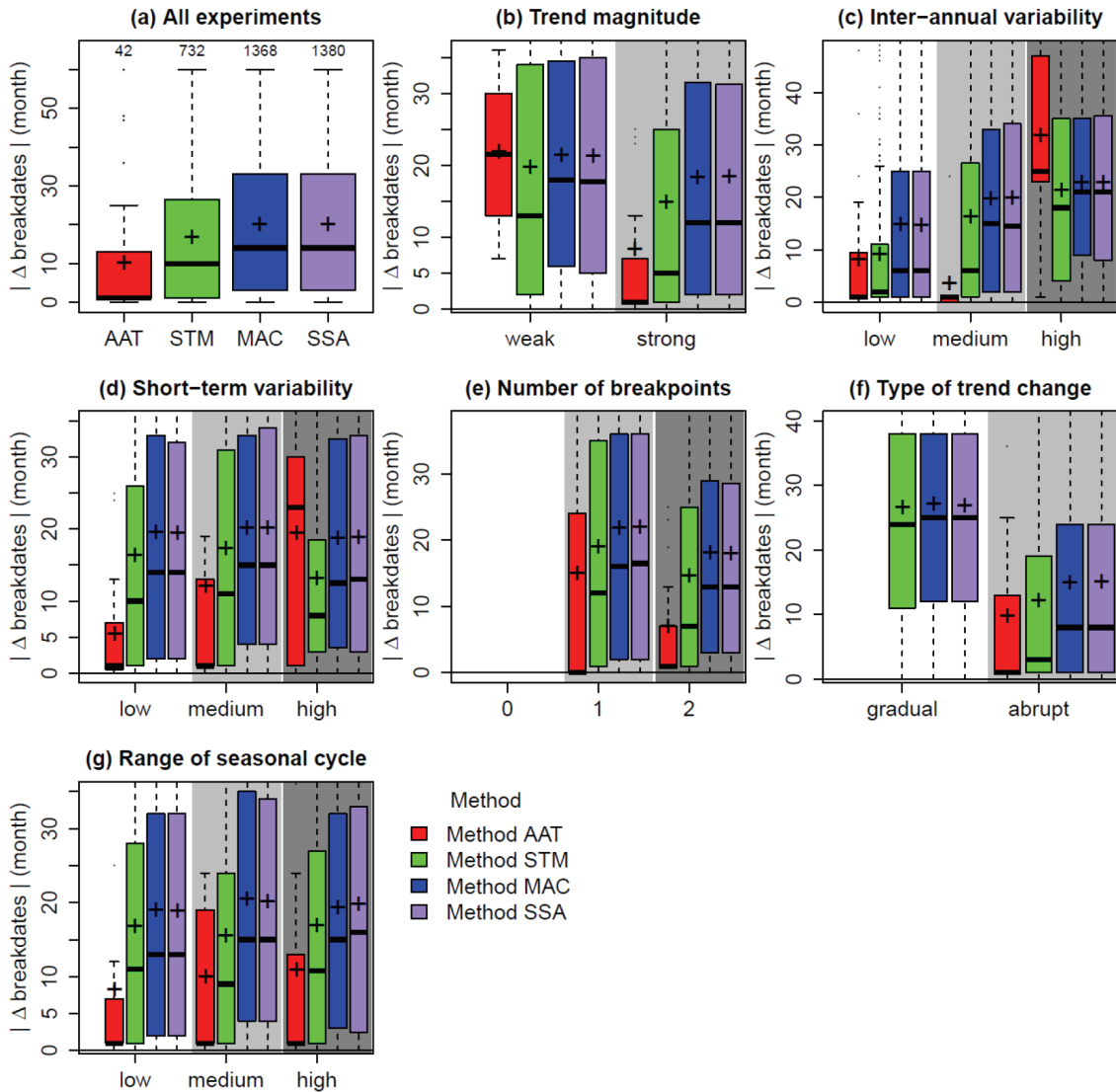


Figure 2.5: Distribution of the temporal absolute difference between real and estimated breakpoints.

(a) Performance of the methods in all experiments. (b) Grouped by trend magnitude. (c) Grouped by inter-annual variability. (d) Grouped by short-term variability. (e) Grouped by the real number of breakpoints. (f) Grouped by the type of trend change. (g) Grouped by the range of the seasonal cycle. + denotes the mean of the distribution. The difference is only based on detected breakpoints. As method AAT detected fewer breakpoints, it has a much smaller sample size ( $n = 42$ ) than the other methods (STM  $n = 732$ , MAC  $n = 1,368$ , SSA  $n = 1,380$ ).

### 2.3.3 Evaluation of estimated trends

In order to evaluate if methods detect the correct trends, the direction and significance of estimated trends were compared against the direction and significance of real trends in a time series segment (Figure 2.6). The estimated slopes from method AAT were higher correlated with the real slopes ( $r = 0.74$ ) than the estimated slopes from other methods (method STM  $r = 0.7$ , MAC  $r = 0.62$  and SSA  $r = 0.61$ ). The agreement was lower if only one of the real or estimated trend was significant ( $r = 0.24$ , AAT) or if

neither the real nor the estimated trends were significant ( $r = 0.22$ , AAT). Nevertheless, estimated trend slopes from all methods were highly correlated (up to  $r = 0.98$  between MAC and SSA, scatterplots not shown). Generally, slope estimates from method AAT were less correlated with the other methods while especially methods MAC and SSA were highly correlated. Hence, the annual aggregation approach in method AAT resulted in the most unique slope estimates compared to the other methods.

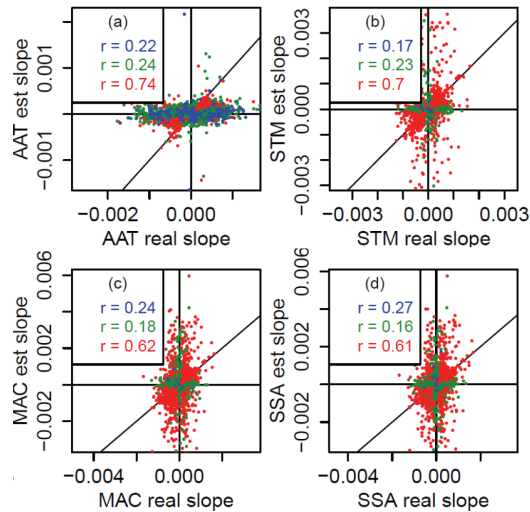


Figure 2.6: Comparison of real and estimated slopes from different methods, all time series segments and all experiments.

Slopes are coloured blue if both real and estimated slopes are not significant, green if only the real or estimated slope was significant and red if both slopes were significant (0.95 significance level).

Confusion matrices were calculated to evaluate the accuracy of trend classes based on the direction and significance of a trend. Method AAT had usually higher class accuracies as well as a higher total accuracy (37.6%) and Kappa coefficient ( $K = 0.25$ ) than other methods (Table 2.1). While method AAT correctly detected 55.24% (47.34%) of significant negative (positive) trends, methods STM, MAC and SSA correctly detected only 47.9% (47.4%), 48.1% (45.15%) and 48.1% (45.9%) of significant negative (positive) trends, respectively. Nevertheless, method AAT detected 3.1% (3.6%), method STM 5.4% (4.6%), method MAC 5.5% (6.1%) and method SSA 5.9% (6.2%) of significant negative (positive) trends as significant positive (negative), i.e., opposite, trends (Table 2.1). Thus, trend slopes from method AAT were higher correlated with real trend slopes and method AAT detected fewer false positive trends and more correct positive trends than methods STM, MAC and SSA.

Table 2.1: Normalized confusion matrices of estimated and real trend classes for each method.

N3: significant negative trend, N2: non-significant negative trend, N1: no trend with negative tendency, P1: no trend with positive tendency, P2: non-significant positive trend, P3: significant positive trend. ToAcc: total normalized accuracy, Kappa: Kappa coefficient.

<b>Method AAT</b>	<b>Real.N3</b>	<b>Real.N2</b>	<b>Real.N1</b>	<b>Real.P1</b>	<b>Real.P2</b>	<b>Real.P3</b>	<b>Sum</b>
Est.N3	55.24	11.18	15.57	8.44	5.95	3.62	100.00
Est.N2	12.48	43.27	26.76	11.11	0.00	6.38	100.00
Est.N1	13.27	14.55	24.55	17.46	17.74	12.42	100.00
Est.P1	10.37	10.57	15.29	24.43	24.85	14.49	100.00
Est.P2	5.54	13.72	11.98	22.01	31.01	15.74	100.00
Est.P3	3.09	6.70	5.85	16.56	20.45	47.34	100.00
Sum	100.00	100.00	100.00	100.00	100.00	100.00	600.00
ToAcc = 37.64, Kappa = 0.25							
<b>Method STM</b>	<b>Real.N3</b>	<b>Real.N2</b>	<b>Real.N1</b>	<b>Real.P1</b>	<b>Real.P2</b>	<b>Real.P3</b>	<b>Sum</b>
Est.N3	47.90	20.68	13.18	7.58	6.07	4.59	100.00
Est.N2	20.58	32.21	14.54	11.05	15.14	6.48	100.00
Est.N1	14.60	18.92	22.68	14.79	18.47	10.54	100.00
Est.P1	10.37	11.09	20.22	25.48	17.20	15.65	100.00
Est.P2	1.15	8.16	19.91	25.21	30.22	15.35	100.00
Est.P3	5.41	8.94	9.47	15.89	12.89	47.39	100.00
Sum	100.00	100.00	100.00	100.00	100.00	100.00	600.00
ToAcc = 34.31, Kappa = 0.21							
<b>Method MAC</b>	<b>Real.N3</b>	<b>Real.N2</b>	<b>Real.N1</b>	<b>Real.P1</b>	<b>Real.P2</b>	<b>Real.P3</b>	<b>Sum</b>
Est.N3	48.08	22.05	11.81	7.56	4.37	6.13	100.00
Est.N2	13.24	29.18	14.06	10.84	26.98	5.69	100.00
Est.N1	15.15	19.12	27.35	14.75	10.87	12.76	100.00
Est.P1	10.91	16.97	15.94	25.79	18.33	12.06	100.00
Est.P2	7.14	4.45	22.71	26.33	21.16	18.22	100.00
Est.P3	5.48	8.23	8.13	14.73	18.29	45.15	100.00
Sum	100.00	100.00	100.00	100.00	100.00	100.00	600.00
ToAcc = 32.79, Kappa = 0.19							
<b>Method SSA</b>	<b>Real.N3</b>	<b>Real.N2</b>	<b>Real.N1</b>	<b>Real.P1</b>	<b>Real.P2</b>	<b>Real.P3</b>	<b>Sum</b>
Est.N3	48.08	17.79	14.94	6.76	6.21	6.22	100.00
Est.N2	9.07	37.14	19.66	11.88	18.57	3.69	100.00
Est.N1	15.20	20.16	24.72	14.54	14.08	11.31	100.00
Est.P1	13.80	9.59	13.99	24.52	24.76	13.34	100.00
Est.P2	7.88	6.19	18.57	25.12	22.70	19.55	100.00
Est.P3	5.98	9.13	8.12	17.17	13.69	45.90	100.00
Sum	100.00	100.00	100.00	100.00	100.00	100.00	600.00
ToAcc = 33.84, Kappa = 0.21							

#### 2.3.4 Effects on the overall performance of the methods

To quantify the overall error of breakpoint and trend detection, the root mean square error (RMSE) between the estimated and real trend component was computed. The distribution of RMSE for the different experimental factors and methods is shown in Figure 2.7. Overall, method AAT and STM performed better than methods MAC and SSA. The error increased with increasing inter-annual variability for all methods (Figure 2.7 c). For all methods the error slightly increased with increasing short-term variability (Figure 2.7 d). All methods had higher errors in time series with breakpoints than in time series without breakpoints (Figure 2.7 e). The range of the seasonal cycle

did not affect the performance of the different methods (Figure 2.7 f). The error was larger in time series with strong trends and abrupt changes than in time series with weak trends and gradual changes (Figure 2.7 b, f). The higher error under strong trends was a result of the worse timing of breakpoints under these conditions (Figure 2.7 b). To quantify the relative effects of the correct timing of breakpoints and inter-annual variability on the error in trend estimation, an analysis-of-variance was computed in the next step.

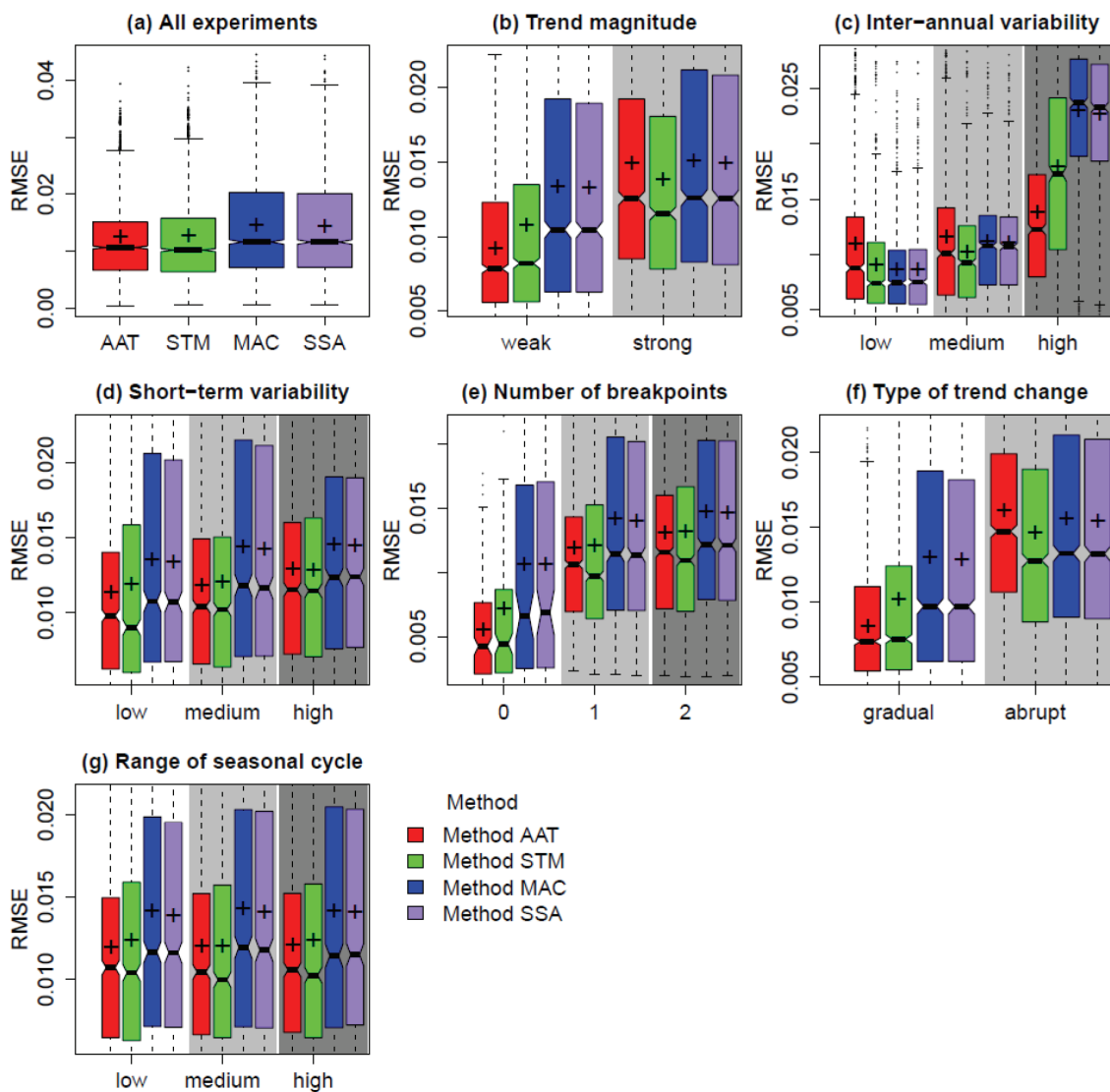


Figure 2.7: Distribution of the root mean square error (RMSE) between real and estimated trend component for the different methods from all experiments.

(a) Performance of the methods in all experiments. (b) Grouped by trend magnitude. (c) Grouped by inter-annual variability. (d) Grouped by short-term variability. (e) Grouped by the real number of breakpoints. (f) Grouped by the type of trend change. (g) Grouped by the range of the seasonal cycle.



The contribution of different factors to the error between estimated and real trend component was analysed by an analysis-of-variance (Table 2.2). All considered factors explained a significant proportion of the error. Inter-annual variability explained the largest part of the error distribution (30.7%). Additionally, the type of trend change, the interaction between inter-annual variability and applied method and the trend magnitude explained large parts of the error. This is illustrated by the similar error of all methods under low and medium inter-annual variability while for methods MAC and SSA the error increased strongly under high inter-annual variability (Figure 2.7 c). The seasonal range had in general only a small contribution to the overall error (0.002%). In summary, the inter-annual variability of the time series was the most important factor for the error of breakpoint and trend estimation.

Table 2.2: Analysis of variance table for the RMSE between real trend and estimated trend. IAV and STV denote inter-annual and short-term variability, respectively. NBP is the number of breakpoints.

Factor	Df	Sum Sq	Mean Sq	F value	P (>F)	Sum Sq / Total Sq (%)
IAV	2	0.2136	0.1068	4096.7	<2.2e-16	30.73
Type of change	1	0.0511	0.0511	1959.3	<2.2e-16	7.35
IAV * Method	6	0.0469	0.0078	299.8	<2.2e-16	6.75
Trend magnitude	1	0.0252	0.0252	966.5	<2.2e-16	3.62
IAV * STV	4	0.0153	0.0038	146.8	<2.2e-16	2.20
Type of change * Method	3	0.0121	0.0040	154.6	<2.2e-16	1.74
Method	3	0.0108	0.0036	137.8	<2.2e-16	1.55
Trend magnitude * Method	3	0.0073	0.0024	93.9	<2.2e-16	1.06
NBP	2	0.0072	0.0036	137.6	<2.2e-16	1.03
STV * Type of change	2	0.0061	0.0030	116.5	<2.2e-16	0.87
STV	2	0.0024	0.0012	46.1	<2.2e-16	0.35
Trend magnitude * Type of change	1	0.0022	0.0022	86.3	<2.2e-16	0.32
Trend magnitude * NBP	2	0.0022	0.0011	41.7	<2.2e-16	0.31
Type of change * NBP	1	0.0022	0.0022	83.4	<2.2e-16	0.31
Trend magnitude * STV	2	0.0022	0.0011	41.3	<2.2e-16	0.31
IAV * Type of change	2	0.0005	0.0003	10.4	2.962E-05	0.08
Seasonality * NBP	4	0.0005	0.0001	5.0	4.945E-04	0.08
STV * NBP	4	0.0005	0.0001	4.5	1.238E-03	0.07
Trend magnitude * IAV	2	0.0004	0.0002	7.4	0.001	0.06
STV * Method	6	0.0003	0.0001	2.1	4.948E-02	0.05
Trend magnitude * Seasonality	2	0.0002	0.0001	4.6	0.010	0.03
Seasonality * Type of change	2	0.0002	0.0001	4.2	1.526E-02	0.03
IAV * NBP	4	0.0002	0.0001	2.0	8.910E-02	0.03
Seasonality	2	0.0000	0.0000	0.2	0.799	0.00
Residuals	10,952	0.2855	0.0000			41.07

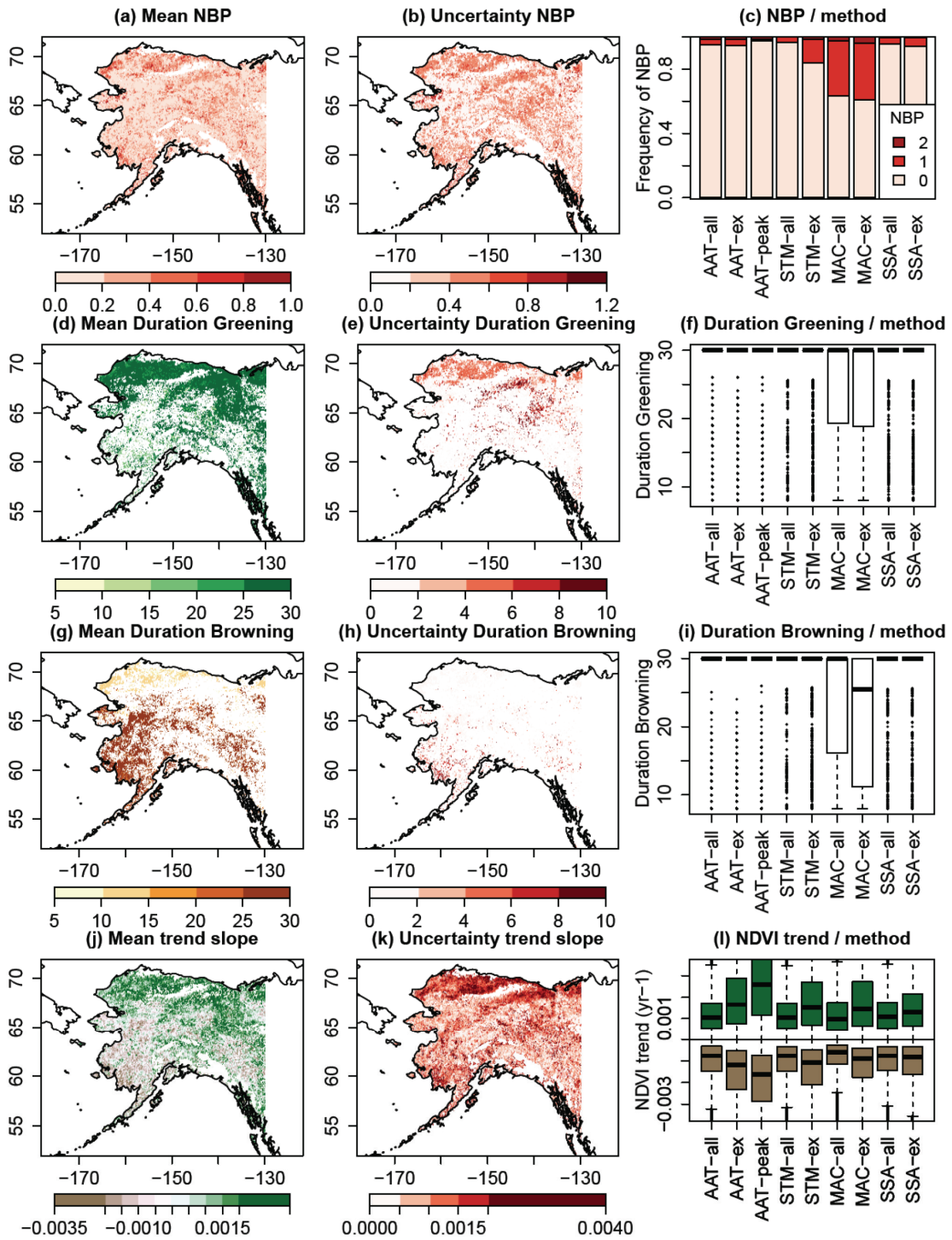


Figure 2.8: Ensemble of breakpoint and trend estimates from all methods in Alaska. AAT, STM, MAC and SSA are the four applied trend methods. ‘all’ indicates that all NDVI values were used (i.e., including interpolated and snow-affected observations). ‘ex’ snow-affected values were excluded from trend analysis. ‘peak’ trend was computed only on annual peak NDVI. (a) Mean number of detected breakpoints from all methods. (b) Uncertainty of the number of detected breakpoints (standard deviation of number of breakpoints from all methods). (c) Number of detected breakpoints grouped by method. (d) mean duration of greening trends (years) with associated uncertainties (e) and distribution of greening duration per each method (f). (g) mean duration of browning trends (years) with associated uncertainties (h) and distribution of browning duration per each method (i). (j) Multi-method mean trend slope ( $\Delta\text{NDVI}/\text{year}$ ) with associated uncertainties (k) and distributions per each method (l).

### 2.3.5 Multi-method ensemble of breakpoint and trend estimates in Alaska

To compare breakpoint and trend estimates of the different methods under real conditions, all methods were applied to GIMMS NDVI3g time series of Alaska and the number and timing of breakpoints as well as the duration of significant greening and browning trends were calculated (Figure 2.8). Breakpoints were detected by multiple methods in northern Tundra regions, in some parts of eastern boreal Alaska and in south-western Alaska (Figure 2.8 a). Method MAC detected in 40 % of all land grid cells one or two breakpoints if snow-affected values were excluded from the analysis (MAC-ex) (Figure 2.8 c). All other methods detected less breakpoints. Especially the seasonal methods detected more breakpoints if snow-affected observation were excluded from the analysis. The lowest number of breakpoints was detected by method AAT applied to the annual peak NDVI (AAT-peak). Despite method SSA, all other methods detected breakpoints around the year 2000 with following browning trends in the northern Tundra. The detection of breakpoints controls the multi-method average duration of significant greening and browning trends. The duration of greening ranges between 20 and up to 30 years with a standard deviation between two to six years in most tundra regions. In some north-eastern parts of the study region significant greening trends last between 20 and 30 years as well but with higher uncertainties of six to 10 years (Figure 2.8 d-e). Browning NDVI trends between 20 and 30 years occurred in some boreal regions of central Alaska and in south-western Alaska usually with small uncertainties of up to four years (Figure 2.8 g-h). The multi-method average NDVI trend slope demonstrates that greening NDVI trends were more spatially and temporally prevalent and of higher magnitude than browning trends (Figure 2.8 j-l). Greening occurred mostly in Tundra regions while browning occurred only spotted in the boreal forest. Nevertheless, greening trends occurred in some parts of the central Alaskan boreal forests too but were associated with higher uncertainties because some methods detected breakpoints with greening NDVI trends while other methods detected no significant greening trends in these regions. All methods detected stronger greening and browning trends if snow-affected values were excluded from the analysis or only peak NDVI was used and weaker trends if all values were used (Figure 2.8 l). In short, methods agreed in the estimation of the major 30 year greening and browning trends, while they had larger differences in regions where breakpoints were detected

by some methods and thus trend changes are likely. Nevertheless, the treatment of snow-affected NDVI values and thus the inherent inter-annual variability of the NDVI time series caused larger differences in NDVI trend estimates than the applied trend estimation method.

## **2.4 Discussion**

### **2.4.1 Effect of temporal resolution on method performance**

The correct timing of breakpoints does not depend on the temporal resolution of the time series but on how full-temporal resolution methods deal with seasonality. Two types of temporal resolution of NDVI time series were explored in this study: Method AAT makes use of a low temporal resolution based on annual aggregated data and methods STM, MAC and SSA were using a monthly temporal resolution of the time series. A more accurate detection of breakpoints was expected if a method uses the full temporal resolution than annual aggregated data. Nevertheless, the annual aggregation approach (method AAT) compared well in the timing of breakpoints like one full-temporal resolution method (method STM) although it largely underestimated the number of breakpoints. Although the other full-temporal resolution methods (methods MAC and SSA) detected more often the right number of breakpoints, they had larger errors in the timing of these breakpoints (Figure 2.5). These results suggest that the estimation of the timing of breakpoints is highly sensitive on how the temporal decomposition methods account for seasonality and inter annual variability. Thus, the lower precision of the annual method limits the correct timing of within-year breakpoints, although it does not reduce the accuracy of this method in comparison to seasonal methods.

The temporal resolution of the time series affects the estimation of the trend significance. The major problem of using annual aggregated data rather than full-resolution time series is the reduction in the number of observations. This involves an underestimation of the significance of the trend (de Beurs and Henebry, 2004b; Verbesselt et al., 2010a). However, this assumed limitation of the annual aggregation approach turned out to be an advantage as it decreases the risk of detecting false positive trends, i.e., a significant positive trend in case of a significant negative real

trend and vice versa. Nevertheless, the application of all methods to real data resulted in similar trend slope estimates from different methods. Similar trend slopes with a low multi-method uncertainty were found in regions where methods usually did not detect breakpoints (Figure 2.8 a ,j). Also previous studies reported similarities in trend estimates from annual-aggregated and full-temporal resolution approaches as well (Mahecha et al., 2010a). Hence, the detection of significant trends depends rather on the estimated breakpoints and thus on time series length than on the temporal resolution of the time series that was used by the trend detection method.

#### **2.4.2 Effect of trend changes on method performance**

The performance of the four methods to estimate trends is lower in time series with breakpoints. All methods estimate trends with a lower error in case of simple time series with no breakpoints or in case of gradual than abrupt changes (Figure 2.7 e, f). All methods present a prevalent tendency to underestimate the number of breakpoints. Method AAT underestimates the number of breakpoints in almost all cases, being the most conservative approach for estimating breakpoints, which ultimately resulted in the best overall performance for trend estimation (Figure 2.7 a). On the other hand, methods MAC and SSA tended to detect breakpoints in time series without breakpoints (Figure 2.4 e). In complex time series with one or multiple breakpoints or abrupt trend changes the error of the estimated underlying trend component is higher for all methods (Figure 2.7 e-f). This higher error is caused by the estimated timing of the breakpoint. A larger difference between the estimated and the real breakpoint causes a higher difference between estimated and real trend slope in the time series segment before or after the breakpoint and thus a higher error in the estimated trend component. The seasonal-adjusted approaches (methods MAC and SSA) have a worse overall performance than the season-trend model approach (method STM) (Figure 2.7 a). Although methods MAC and SSA detected in more cases the right number of breakpoints they had a significant number of false positive detected breakpoints (Figure 2.4 a). The overestimation and/or the strong mismatch in the timing of the breakpoints cause a false detection of the underlying trend or even a detection of opposite trends, especially under high inter-annual variability (Figure 2.5 c). Taken together, in order to detect the correct trends, the underestimation of breakpoints

results in better trend estimates, and better estimates of the timing of major breakpoints are more important than the detection of the right number of breakpoints.

### **2.4.3 Effect of inter-annual variability on method performance**

The capability to estimate trends and breakpoints depends mostly on the robustness of a method against inter-annual variability (Figure 2.7, Table 2.2). Inter-annual variability increased the timing error of breakpoints (Figure 2.5 c). The error of the estimated trend component increased for all methods with increasing inter-annual variability. Nevertheless, the error increase with increasing inter-annual variability is much stronger for methods STM, MAC and SSA than for method AAT (Figure 2.7 c). Thus, the annual aggregation method is a relatively robust approach against inter-annual variability for estimating trends and trend changes. This effect is also evident from the application of the methods on real NDVI time series: The inter-annual variability of NDVI time series in Alaska is 1.6 times higher if snow-affected observations are excluded (figure not shown). This inter-annual variability caused a detection of a higher number of breakpoints if snow-affected observation were not used to compute breakpoints in all methods (Figure 2.8 c). The use or non-use of snow-affected NDVI observations for trend analysis and the associated inter-annual variability of the NDVI time series caused larger differences in breakpoint estimates and trend slopes than the choice of the trend method (Figure 2.8 l).

However, we have to be cautionary in the assessment of the effect of inter-annual variability on the method performance: We assumed that the inter-annual variability as well as the short-term variability are independent, i.e., temporally uncorrelated processes. Real world observations may depict long-range correlations that can be expected to induce year-to-year dependencies in the inter-annual variability. This would imply that the separation of trend and inter-annual variability is not straightforward. However, an investigation on effects of this kind is beyond the scope of our study but we assume that additional uncertainties would affect the timing of changes in trends in any methodological approach and attribution study.

#### **2.4.4 Plausibility of trend and breakpoint estimates in Alaska**

Alaska is of special interest for the analysis of trend change detection methods because previous studies reported greening NDVI trends in tundra ecosystems of the Alaska North Slope as well as browning trends in the Alaska boreal forests (Beck and Goetz, 2011; Goetz et al., 2005; de Jong et al., 2011a; Neigh et al., 2008). On the other hand, it is not clear if these browning trends are monotonic trends as also different number of trend changes and a considerable amount of years with greening trends were detected (de Jong et al., 2013b). No trend changes or browning trends have been reported in previous studies for the Alaskan tundra (de Jong et al., 2013b). We detected breakpoints in the Alaskan tundra around the year 2000 with following browning trends across different methods and regardless if snow-affected NDVI observations or only peak NDVI were used for trend analysis or not (Figure 2.8 g–i). That means that these breakpoints with following browning in the Alaskan tundra are not artefacts due to snow or cloud contaminations or other radiative surface changes outside the peak-growing season. This browning is observable due to the extended length of the GIMMS NDVI3g dataset until 2011 in comparison to previous dataset versions. It needs to be assessed if these greening-to-browning trend changes in the Alaskan tundra are due to regime shifts in climate-ecosystem interaction processes or due to inter-annual variations. Browning trends were detected by all methods in some parts of boreal Alaska. Previous studies reported negative NDVI trends especially in eastern boreal Alaska (Parent and Verbyla, 2010). These browning trends are associated with wildfires (Goetz et al., 2005), occur mainly in evergreen needle-leaf forests (Beck and Goetz, 2011) and are related to temperature-induced drought stress and insect disturbances (Parent and Verbyla, 2010; Verbyla, 2011). Browning trends in western boreal Alaska based on a previous version of the GIMMS NDVI dataset are more doubtful because such GIMMS NDVI had only a weak agreement with MODIS NDVI in this often cloud-affected region (Parent and Verbyla, 2010). In contrast to previous results of de Jong et al. (2011b), we detected fewer trend changes in boreal Alaska and more monotone browning trends over this 30 year period.

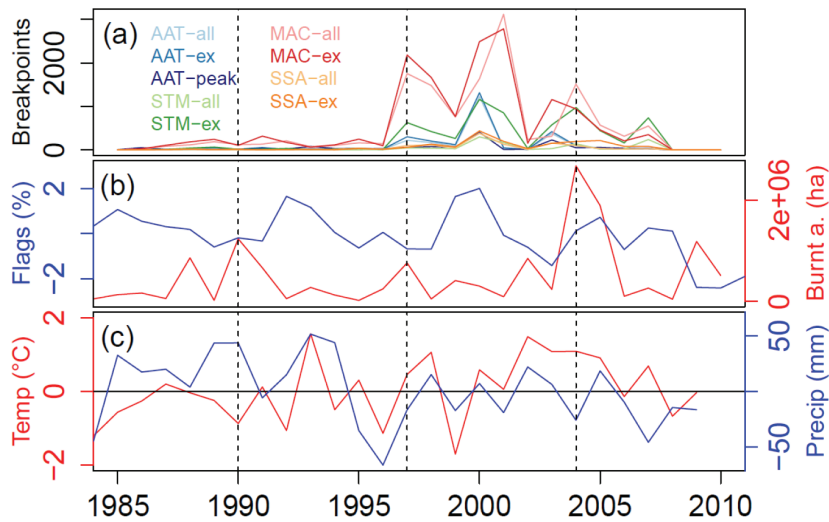


Figure 2.9: Comparison of detected breakpoints with temporal fire and climate patterns in Alaska.

(a) Time series of the total number of detected breakpoints per year for each method. (b) Total annual burnt area and annual anomaly of flagged GIMMS NDVI3g pixels with reduced quality. (c) Annual temperature and precipitation anomalies (baseline 1982–2009) averaged for Alaska.

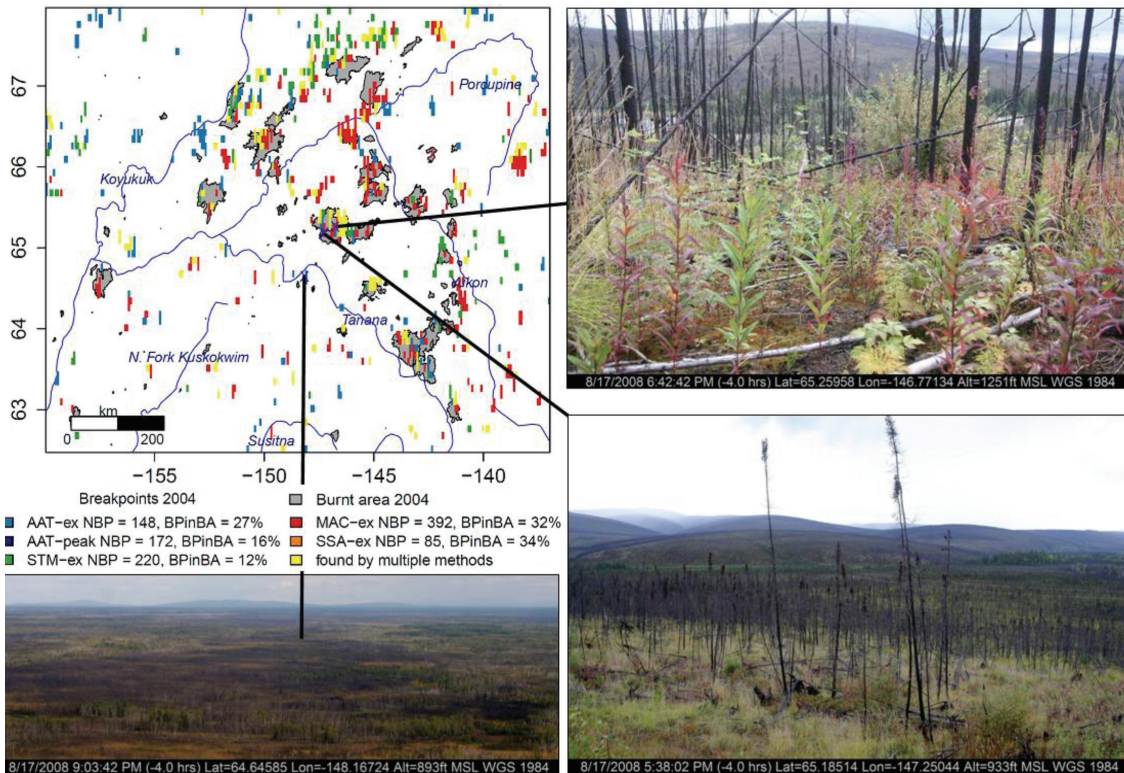


Figure 2.10: Comparison of detected breakpoints of the year 2004 from four different methods with 2004 burnt areas and *in situ* photos (taken in 2008).

NBP is the total number of breakpoints that was detected in 2004 in this region. BPinBA denotes the percentage of breakpoints that was found inside a burnt area polygon. For methods AAT-ex and AAT-peak breakpoints for both 2003 and 2004 are shown to compensate for the lower breakpoint timing precision of these methods. All breakpoints that were found by AAT-peak were found at least also by one other method.



To evaluate the plausibility of detected breakpoints in NDVI time series of Alaska, the temporal dynamics and spatial patterns of breakpoints were compared against quality flags of the GIMMS NDVI3g dataset, temperature and precipitation anomalies as well as burnt areas from the Alaskan Large Fire Database (Figure 2.9, Figure 2.10). We cannot fully make sure that all detected breakpoints are only due to climate or environmental changes but might be caused by sensor changes in the GIMMS NDVI3g dataset. Sensor changes took place in 1985, 1988, 1994, 1995, 2000 and 2004 (Tucker et al., 2005). Although most breakpoints were detected in the year 2000 which had the highest number of flagged NDVI values with reduced quality and where a sensor change took place (Figure 2.9), previous studies have shown that trends computed on GIMMS NDVI data are not affected by sensor changes (Kaufmann et al., 2000; Tucker et al., 2005). Clearly, some breakpoints can be related to drought periods or fire events. Some larger fire events occurred in 1988, 1990, 1997, 2002 and 2004 and methods detected breakpoints in NDVI trends during or one year after these fire events (Figure 2.9 a-b). Methods AAT, STM and MAC detected breakpoints in 1997 and 1998 in central and south-western Alaska that can be related to patterns of negative precipitation anomalies and fire activity. Alaskan fire activity is caused by droughts that are related to large-scale circulation patterns (Macias Fauria and Johnson, 2006). For example, El Nino causes drought and fire weather conditions in interior Alaska (Hess et al., 2001; Macias Fauria and Johnson, 2006). Thus it is probable, that detected breakpoints in 1997/1998 are a result of decreased NDVI signals because of drought effects after the 1997 El Nino event (McPhaden, 1999). The largest total burnt area in Alaska occurred in 2004 under low precipitation and high temperature conditions (Figure 2.9 b-c). The detection of breakpoints in NDVI time series in 2004 agreed with the spatial distribution of wildfires and in situ observations. Large conifer forests were burned in 2004 and were replaced by low broadleaved shrubs (dwarf birch and aspen) and grasses during post-fire succession which results in a structural change in NDVI time series (Figure 2.10). Especially methods AAT and MAC detected the highest number of breakpoints exactly at burn scar locations (AAT-ex n = 40, MAC n = 124) while methods STM and SSA found less breakpoints at burnt areas (STM n = 27 and SSA n = 29). Nevertheless, only a few fire-related breakpoints were detected by multiple methods and especially methods STM and MAC detected many breakpoints

outside burnt areas. As method MAC had a high number of false positive detected breakpoints (Figure 2.4), some of these breakpoints might be false positives too. However, we cannot assess if they were caused by other environmental processes. In short, detected breakpoints can be related to environmental conditions like fire events or drought periods but need to be cross-checked against quality flags of the GIMMS NDVI3g dataset and additional environmental datasets to exclude false positive detected breakpoints.

#### **2.4.5 Practical recommendations**

Based on our results we want to summarize advantages and limitations of the methods and give recommendations for practical applications of trend and breakpoint detection methods on long-term NDVI time series. All tested methods offer advantages but involve also different limitations for trend and breakpoint estimation. Removing a mean annual cycle from seasonal time series (method MAC) in order to calculate trends is an often applied method. However, even if this trend analysis on seasonal-adjusted time series had the highest number of correct detected breakpoints, it involves also the highest number of false positive detected breakpoints and had a low overall performance for trend and breakpoint detection. The method that removes a modulated annual cycle as detected by singular spectrum analysis (method SSA) allows distinguishing and quantifying changes that are caused by changes in seasonality or caused by the long-term NDVI trend. Although this method resulted in a high proportion of correct detected breakpoints, it involves a high number of false positive detected breakpoints as method MAC. Additionally, the de-seasonalisation by a modulated annual cycle can remove the inter-annual variability that is related to trends and results in a low overall performance of this method. Trend and breakpoint estimation based on a season-trend model (method STM) quantifies trends while taking into account the seasonality and noise of the NDVI time series. Thus, it allows detecting, distinguishing and quantifying changes in the phenological cycle as well as in the long-term NDVI trend (Verbesselt et al., 2010a, 2010b, 2012). Although this method under-estimated the number of breakpoints and had 5% false positive detected breakpoints, it had a good overall performance. Trend analysis on annual aggregated values (method AAT) highly underestimated the number of breakpoints but had the

best overall performance of all methods considering the estimation of trends. The annual aggregation method had the lowest number of wrong detected breakpoints and detected real breakpoints were found with the highest accuracy of all methods (Figure 2.5). While this method reduces the temporal resolution and has a lower precision for the timing breakpoints, it offers the potential to calculate trends on specific properties of NDVI time series like annual or growing season means, as well as peak NDVI or seasonal amplitudes. The specific properties of NDVI time series in high-latitude regions (short vegetation period, snow cover, often cloud affected) limit the applicability of the seasonal methods and ultimately suggest using method AAT on peak NDVI to exclude observations affected by distortions.

Based on different advantages and limitations of all methods, we recommend using method AAT on mean growing season or peak NDVI for regions where the time series are often affected by distortions. If the seasonal NDVI time series values outside the peak NDVI period are credible, the calculation of a multi-method ensemble based on full time series could help to detect robust trends and breakpoints assuming that the agreement of multiple methods is more reliable than a single method. The later approach allows not only to detect trends but also to quantify the uncertainty of the trend estimate based on the choice of the trend method. Breakpoints can be considered as more robust if they were detected by multiple methods. Nevertheless the environmental plausibility of detected breakpoints needs to be assessed. Breakpoints with abrupt changes, i.e., with higher magnitudes, were detected more accurately than breakpoints with gradual changes, i.e., with low magnitudes. On the other hand false detected abrupt changes caused a low overall performance of all trend methods. Thus, we recommend to check the magnitude of changes at breakpoints and to relate these breakpoints in a driver-oriented framework (de Jong et al., 2013b) to potential causes of changes like land cover changes, drought or disturbances as fire or insect infestations.

As the detection of breakpoints causes additional uncertainties in trend estimates, the purpose of the breakpoint detection in analysis of NDVI trends needs to be clearly defined: Are trends or trend changes (breakpoints) in the focus of a study? Although the detection of breakpoints offer the potential to detect disturbances in NDVI time series, trend changes and trends in sub-segments of the NDVI time series, a false

detection of non-existing breakpoints can result in wrong or even opposite NDVI trend estimates.

Inter-annual variability is the most important factor for the performance of methods to detect trends and breakpoints in NDVI time series. Main sources for inter-annual variability in NDVI time series are (1) contaminants like insufficient pre-processed data or insufficient harmonized multi-sensor observations, (2) meteorological distortions like clouds, dust, aerosols or snow cover and (3) environmental processes like climate variability, disturbances and land cover changes with associated changes in ecosystem structure and productivity. Usually, only the later source of inter-annual variability is of interest in NDVI time series analyses. Users of long-term NDVI datasets rely on the pre-processing and harmonization of multi-sensor observations performed in all conscience by dataset providers. Meteorological distortions can be excluded from analyses by excluding NDVI observations that are flagged as snow or poor quality; by aggregating the bi-monthly GIMMS NDVI3g dataset to monthly temporal resolution; or by analysing only annual peak NDVI observations. As snow cover and clouds have low NDVI values, an extended snow cover can likely cause a detection of weaker NDVI greening or even the detection of browning trends. Thus, the use or non-use of snow-affected and low quality NDVI observations directly affects the inter-annual variability of the NDVI time series and such NDVI values should be excluded from trend and breakpoint analyses.

## **2.5 Conclusions**

We demonstrated that increasing inter-annual variability in Normalized Difference Vegetation Index (NDVI) time series decreases the performance of methods to detect trends and trend changes in long-term NDVI time series. Trend estimation based on annual aggregated NDVI time series and the season-trend method had good overall performances. Hence, in order to detect trend changes in NDVI time series with higher precision and accuracy, one needs to improve methods that work on the full temporal resolution time series regarding the robustness against inter-annual variability. Inter-annual variability of NDVI time series can be caused by artefacts from the harmonization of a dataset from different sensors, meteorological distortions like

clouds or snow and environmental processes such as climate patterns or disturbances. As a consequence, snow-affected NDVI observation or observations with poor quality need to be excluded from trend and breakpoints analyses as the performance of trend and breakpoint estimation methods decreases with increasing inter-annual variability. Methods can detect wrong or even opposite NDVI trends if they detect breakpoints in time series that have no breakpoints. Nevertheless, the detection of breakpoints offers the potential to detect trend changes or disturbances in NDVI time series.

We evaluated for the first time different methods to detect trends and trend changes in newly available 30 year GIMMS NDVI3g time series. Future studies of trends and breakpoints in long-term NDVI time series should assess the plausibility of detected breakpoints against multi-method ensemble estimates of breakpoints, quality flags of the NDVI time series and further environmental data streams, in order to prevent a detection of wrong NDVI trends.

## **Acknowledgements**

We thank Compton Tucker, Jorge Pinzon, Ranga Myneni and the GIMMS group for producing and providing the GIMMS NDVI3g dataset. We thank Martin Jung for his comments on this work. We are very thankful to the comments from six anonymous reviewers that helped improving the quality of the manuscript.

We thank the community of the R statistical software for providing a wealth of functionality. The methods described in this article are available as R package under the GNU General Public License on <http://greenbrown.r-forge.r-project.org/>.

M.F. received funding from the Max Planck Institute for Biogeochemistry and from the European Commission's 7th Framework Programme project CARBONES (grant agreement 242316). N.C. was supported by the European Commission's 7th Framework Program project Carbo-Extreme (grant agreement 226701). J.V. was supported by a Marie-Curie IRG grant of the European Commission's 7th Framework Program (grant agreement 268423). M.M. was supported by the European Commission's 7th Framework Program project GEOCARBON (grant agreement 283080). C.N. was supported by NASA's Terrestrial Ecology and Carbon Cycle Science

Programs (grant agreement NNH07ZDA001N-CARBON). M.F conducted this work under the International Max Planck Research School for Global Biogeochemical Cycles.



# 3 Identifying environmental controls on vegetation greenness phenology through model-data integration

---

This chapter was originally published as:

Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

The Supplement of this chapter can be found in chapter 7.

Biogeosciences, 11, 7025–7050, 2014  
www.biogeosciences.net/11/7025/2014/  
doi:10.5194/bg-11-7025-2014  
© Author(s) 2014. CC Attribution 3.0 License.



## Identifying environmental controls on vegetation greenness phenology through model–data integration

M. Forkel<sup>1</sup>, N. Carvalhais<sup>1,2</sup>, S. Schaphoff<sup>3</sup>, W. v. Bloh<sup>3</sup>, M. Migliavacca<sup>1</sup>, M. Thurner<sup>1,4</sup>, and K. Thonicke<sup>3</sup>

<sup>1</sup>Max-Planck-Institute for Biogeochemistry, Department for Biogeochemical Integration, Hans-Knöll-Str. 10, 07745 Jena, Germany

<sup>2</sup>Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, 2829-516, Caparica, Portugal

<sup>3</sup>Potsdam Institute for Climate Impact Research, Earth System Analysis, Telegraphenberg A31, 14473 Potsdam, Germany

<sup>4</sup>Stockholm University, Department of Applied Environmental Science and Bolin Centre for Climate Research, Svante Arrhenius väg 8, 10691 Stockholm, Sweden

*Correspondence to:* M. Forkel (mforkel@bgc-jena.mpg.de)

Received: 28 May 2014 – Published in *Biogeosciences Discuss.*: 17 July 2014

Revised: 10 October 2014 – Accepted: 21 October 2014 – Published: 11 December 2014



### 3 Identifying environmental controls on vegetation greenness phenology through model-data integration

#### Abstract

Existing dynamic global vegetation models (DGVMs) have a limited ability in reproducing phenology and decadal dynamics of vegetation greenness as observed by satellites. These limitations in reproducing observations reflect a poor understanding and description of the environmental controls on phenology, which strongly influence the ability to simulate longer-term vegetation dynamics, e.g. carbon allocation. Combining DGVMs with observational data sets can potentially help to revise current modelling approaches and thus enhance the understanding of processes that control seasonal to long-term vegetation greenness dynamics. Here we implemented a new phenology model within the LPJmL (Lund Potsdam Jena managed lands) DGVM and integrated several observational data sets to improve the ability of the model in reproducing satellite derived time series of vegetation greenness. Specifically, we optimized LPJmL parameters against observational time series of the fraction of absorbed photosynthetic active radiation (FAPAR), albedo and gross primary production to identify the main environmental controls for seasonal vegetation greenness dynamics. We demonstrated that LPJmL with new phenology and optimized parameters better reproduces seasonality, inter-annual variability and trends of vegetation greenness. Our results indicate that soil water availability is an important control on vegetation phenology not only in water-limited biomes but also in boreal forests and the Arctic tundra. Whereas water availability controls phenology in water-limited ecosystems during the entire growing season, water availability co-modulates jointly with temperature the beginning of the growing season in boreal and Arctic regions. Additionally, water availability contributes to better explain decadal greening trends in the Sahel and browning trends in boreal forests. These results emphasize the importance of considering water availability in a new generation of phenology modules in DGVMs in order to correctly reproduce observed seasonal-to-decadal dynamics of vegetation greenness.

### 3.1 Introduction

The greenness of the terrestrial vegetation is directly linked to plant productivity, surface roughness and albedo and thus affects the climate system (Richardson et al., 2013). Vegetation greenness can be quantified from satellite observations for example as Normalized Difference Vegetation Index (NDVI) (Tucker, 1979). NDVI is a remotely sensed proxy for structural plant properties like leaf area index (LAI) (Turner et al., 1999), green leaf biomass (Gamon et al., 1995) and plant productivity. In particular, NDVI of green vegetation has a linear relationship with the fraction of absorbed photosynthetic active radiation (FAPAR) (Fensholt et al., 2004; Gamon et al., 1995; Myneni et al., 1995, 1997b; Myneni and Williams, 1994). Satellite-derived FAPAR estimates are often used to estimate terrestrial photosynthesis (Beer et al., 2010; Jung et al., 2008, 2011; Potter et al., 1999). Decadal satellite observations of NDVI demonstrate widespread positive trends (“greening”) especially in the high-latitude regions (Lucht et al., 2002; Myneni et al., 1997a; Xu et al., 2013) but also in the Sahel, southern Africa and southern Australia (Fensholt and Proud, 2012; de Jong et al., 2011b, 2013b). Surprisingly, these trends are accompanied by negative trends (“browning”) which were observed regionally in parts of the boreal forests of North America and Eurasia, and in parts of eastern Africa and South America (Baird and Verbyla, 2012; Bi et al., 2013; de Jong et al., 2013b). Regionally different causes have been identified for the observed greening and browning trends. The greening of the high latitudes is supposed to be mainly induced by rising air temperatures (Lucht et al., 2002; Myneni et al., 1997a; Xu et al., 2013). Browning trends in subtropical regions were related to changing drought conditions and land use change (Cook and Pau, 2013; van Leeuwen et al., 2013). On the other hand, the environmental controls on the browning of boreal forests have been intensively investigated but no concluding or general explanation has been found so far (Barichivich et al., 2014; Beck et al., 2011; Beck and Goetz, 2011; Bunn et al., 2007; Goetz et al., 2005; Piao et al., 2011; Wang et al., 2011). Trends in vegetation greenness are often related to changes in vegetation phenology like an earlier onset and an associated lengthening of the growing season in mid- and high-latitude regions (Atzberger et al., 2013; Høgda et al., 2001, 2013; Tucker et al., 2001; Zeng et al., 2011). Changes in vegetation greenness are linked to changes in primary production and thus affect atmospheric CO<sub>2</sub> concentrations and the terrestrial carbon

cycle (Barichivich et al., 2013; Keeling et al., 1996; Myneni et al., 1997a). Additionally, vegetation greenness affects the climate system by influencing surface albedo. For example, greening trends in high latitudes are associated with decreasing surface albedo (Urban et al., 2013) which alters the surface radiation budget (Lorantý et al., 2011). This can potentially further contribute to a warming of Arctic regions (Chapin et al., 2005). Thus, satellite observations of vegetation greenness demonstrate the recent interactions and changes between terrestrial vegetation dynamics and the climate system.

Dynamic global vegetation models (DGVM) or generally climate/carbon cycle models are used to analyse and project the response of the terrestrial vegetation to the past, recent and future climate variability (Prentice et al., 2007). DGVMs can be used to explain observed trends in vegetation greenness (Lucht et al., 2002) or to quantify the related terrestrial CO<sub>2</sub> uptake. While most global models simulate an increasing uptake of CO<sub>2</sub> by the terrestrial vegetation under future climate change scenarios, the magnitude of future changes in land carbon uptake largely differs among models (Friedlingstein et al., 2006; Sitch et al., 2008). The spread of land carbon uptake estimates among DGVMs might be partly related to insufficient representations of vegetation phenology and greenness (Richardson et al., 2012). Coupled climate–carbon cycle models and uncoupled DGVMs have been compared against 30-year satellite-derived time series of LAI (Anav et al., 2013; Murray-Tortarolo et al., 2013; Zhu et al., 2013). Models usually overestimate mean annual LAI in all biomes and have a too long growing season because of a delayed season end (Anav et al., 2013; Murray-Tortarolo et al., 2013; Zhu et al., 2013). Additionally, most DGVMs have more positive LAI trends than the satellite-derived LAI product, i.e. they underestimate browning trends in boreal forests while a few DGVMs do not reproduce the general greening of the high latitudes (Murray-Tortarolo et al., 2013). The limitations of DGVMs in reproducing observed LAI or FAPAR time series is mostly related to limited phenology routines that often miss environmental controls on seasonal leaf development (Kelley et al., 2013; Murray-Tortarolo et al., 2013; Richardson et al., 2012). In conclusion, with improved modelling approaches for vegetation phenology and greenness, DGVMs can potentially more accurately reproduce the recent, and project the future response of the terrestrial vegetation to climate variability.

Past studies successfully demonstrated the use of vegetation greenness observations to improve stand-alone phenology models or to optimize phenology and productivity-related parameters in DGVMs. The growing season index (GSI) is an empirical phenology model that is used to estimate seasonal leaf developments (Jolly et al., 2005). Empirical parameters of GSI have been optimized against globally distributed 10-year FAPAR and LAI time series from MODIS to reanalyse climatic drivers for vegetation phenology (Stöckli et al., 2008b, 2011). This optimization resulted in a good representation of temporal FAPAR and LAI dynamics in all major biomes except evergreen tropical forests (Stöckli et al., 2011). Model parameters of the Biome-BGC model were optimized against eddy covariance flux observations and NDVI time series from MODIS for poplar plantations in northern Italy which resulted in a more accurate representation of carbon fluxes and NDVI (Migliavacca et al., 2009). The BETHY-CCDAS model was optimized against FAPAR time series from MERIS for seven eddy covariance sites (Knorr et al., 2010) and later for 170 land grid cells using coarse 8 by 10° spatial resolution (Kaminski et al., 2012). These studies demonstrated the improvements in simulated vegetation phenology by optimizing model parameters against observations of vegetation greenness.

Nevertheless, spatial patterns and temporal dynamics of vegetation greenness were not yet optimized in a DGVM globally at a higher spatial resolution (0.5°) and by using long-term (30 year) satellite-derived time series of vegetation greenness. Newly developed 30 year time series of LAI or FAPAR from the GIMMS3g data set (Global Inventory Modeling and Mapping Studies, third generation of data sets; Zhu et al., 2013) make it possible to improve DGVMs not only based on seasonal cycles of single years (i.e. phenology) but additionally against decadal time series properties including inter-variability and trends. By integrating the GIMMS3g FAPAR data set in a DGVM, we can potentially improve spatial patterns and seasonal to long-term temporal dynamics of vegetation greenness. We use the LPJmL DGVM (Lund–Potsdam–Jena managed lands). Similar to other DGVMs, LPJmL does not accurately reproduce the growing season onset and seasonal amplitude of observed LAI and FAPAR time series, presumably because of a limited phenology model (Kelley et al., 2013; Murray-Tortarolo et al., 2013). Thus integrating long-term observations of FAPAR in the LPJmL DGVM potentially requires the development of an improved phenology scheme.

We aim to improve environmental controls on vegetation phenology and greenness in LPJmL by (1) developing a new phenology module for LPJmL, by (2) optimizing FAPAR, productivity and phenology-related parameters of LPJmL against 30-year satellite-derived time series of FAPAR, against 10-year satellite-derived time series of vegetation albedo and against spatial patterns of mean annual gross primary production (GPP) from a data-oriented estimate and by (3) integrating further data streams into LPJmL to constrain land cover dynamics and disturbance effects on vegetation greenness in diagnostic model simulations. This model-data integration approach for LPJmL (LPJmL-MDI) will be applied to identify the environmental controls on vegetation greenness phenology.

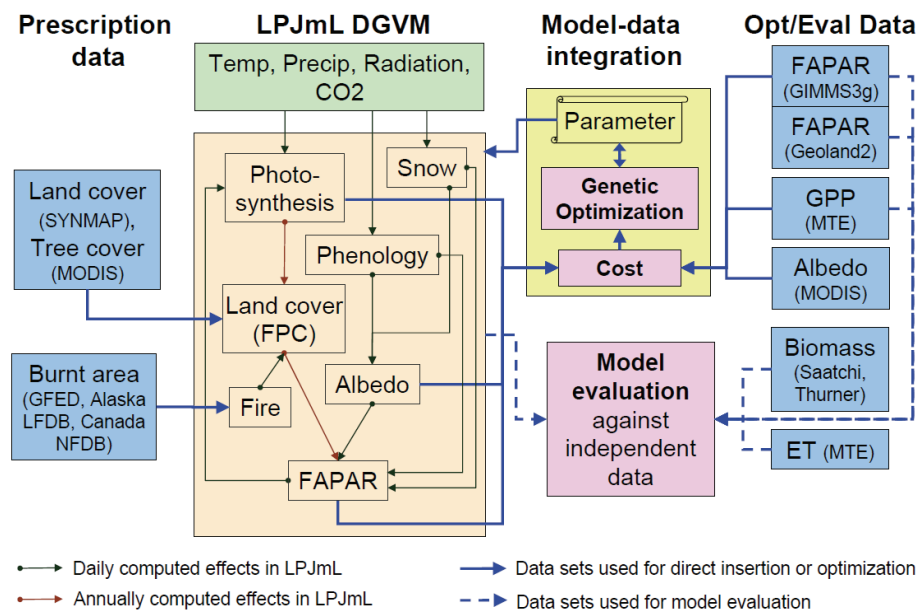


Figure 3.1: Structure of the model–data integration approach for LPJmL (LPJmL-MDI). The LPJmL model structure is highly simplified.

## 3.2 Model, data sets and model–data integration

### 3.2.1 Overview

LPJmL is a dynamic global vegetation model that simulates ecosystem processes such as carbon and water fluxes, carbon allocation in plants and soils, permafrost dynamics, fire spread and behaviour and vegetation establishment and mortality. We used LPJmL version 3.5. This version is based on the original LPJ model (Sitch et al., 2003). The model has been extended for human land use (Bondeau et al., 2007), and agricultural

water use (Rost et al., 2008). It includes a process-oriented fire model (Thonicke et al., 2010), an improved representation of surface albedo and snow coverage (Strengers et al., 2010) and a newly implemented soil hydrology scheme and permafrost module (Schaphoff et al., 2013). This study focuses on the natural vegetation plant functional types (PFTs) (Sitch et al., 2003), i.e. our model developments and optimizations were not applied for crop functional types (CFTs) (Bondeau et al., 2007) because crop phenology is highly driven by human practices.

We developed a model–data integration approach for the LPJmL DGVM (LPJmL-MDI, Figure 3.1). LPJmL-MDI allows us (1) to directly insert land cover, tree cover and burnt area data sets in LPJmL for diagnostic model applications (Sect. 3.2.4.1); (2) to optimize LPJmL model parameters against data sets (here FAPAR, GPP, albedo; Sect. 3.2.4.2); and (3) to evaluate and benchmark LPJmL simulations against observations or observation-based data sets (Sect. 3.2.4.3). Like in a prognostic mode, LPJmL was driven by climate forcing data. Additionally, observed burnt areas were directly inserted into LPJmL to consider observed fire dynamics in diagnostic model applications. For this, we directly replaced the simulated burnt area in the LPJmL-SPITFIRE fire module (Thonicke et al., 2010) by observed burnt areas using the approach of Lehsten et al. (2008). Thus, the timing and location of fire spread is constrained by observations whereas fire effects on vegetation are still simulated by LPJmL-SPITFIRE. We further prescribed observed land cover and tree cover fractions to control for vegetation dynamics in parameter optimization experiments. Observed FAPAR and albedo time series as well as observation-based mean annual spatial patterns of GPP were used in a joint cost function to optimize productivity, phenology, radiation, and albedo-related model parameters using a genetic optimization algorithm.

LPJmL was previously evaluated against site measurements of net carbon ecosystem exchange (Schaphoff et al., 2013; Sitch et al., 2003), atmospheric CO<sub>2</sub> fractions (Sitch et al., 2003), soil moisture (Wagner et al., 2003), evapotranspiration and runoff (Gerten et al., 2004; Schaphoff et al., 2013), fire regimes (Thonicke et al., 2010), and permafrost distribution (Schaphoff et al., 2013). Here we evaluate LPJmL against additional and partly new available global data sets of FAPAR (Baret et al., 2013; Zhu et al., 2013), GPP

and evapotranspiration (ET) (Jung et al., 2011), tree cover (Townshend et al., 2011) and biomass (Carvalhais et al., 2014; Saatchi et al., 2011; Thurner et al., 2014).

### 3.2.2 FAPAR and phenology in the LPJmL DGVM

#### 3.2.2.1 FAPAR

FAPAR is defined as the ratio between the photosynthetic active radiation absorbed by the green canopy (APAR) and the total incident photosynthetic active radiation (PAR). Thus, the total FAPAR of a grid cell is the sum of FAPAR that is distributed among the individual PFTs:

$$FAPAR_{PFT} = \frac{APAR_{PFT}}{PAR} \quad (3.1)$$

$$FAPAR_{gridcell} = \sum_{PFT=1}^{PFT=n} FAPAR_{PFT} \quad (3.2)$$

where  $n$  is the number of established PFTs in a grid cell. The FAPAR of a PFT depends on the annual maximum foliar projective cover (FPC), on the daily snow coverage in the green canopy ( $F_{snow,gv}$ ), green-leaf albedo ( $\beta_{leaf}$ ) and the daily phenology status ( $Phen$ ):

$$FAPAR_{PFT} = FPC_{PFT} \times (Phen_{PFT} - (Phen_{PFT} \times F_{snow,gv,PFT})) \times (1 - \beta_{leaf,PFT}) \quad (3.3)$$

Thus, the temporal dynamic of FAPAR in LPJmL is affected on an annual time step by changes in foliar projective cover ( $FPC_{PFT}$ ) and on daily time steps by changes in phenology ( $Phen_{PFT}$ ) and snow coverage in the green canopy ( $F_{snow,gv,PFT}$ ) (Figure 7.1). This approach extends the previous implementation of Sitch et al. (2003) where FAPAR depends only on FPC and phenology but leaf albedo and snow effects on FAPAR were missing.

$FPC_{PFT}$  expresses the land cover fraction of a PFT. It is the annual maximum fractional green canopy coverage of a PFT and is annually calculated from crown area ( $CA$ ), population density ( $P$ ) and LAI (Sitch et al., 2003):

$$FPC_{PFT} = CA_{PFT} \times P_{PFT} \times (1 - e^{-k_{PFT} \times LAI_{PFT}}) \quad (3.4)$$

The last term expresses the light extinction in the canopy which depends exponentially on LAI and the light extinction coefficient  $k$  of the Lambert-Beer law (Monsi and Saeki, 1953). The parameter  $k$  had a constant value of 0.5 for all PFTs in the original LPJmL formulation (Sitch et al., 2003). We changed  $k$  to a PFT-dependent parameter because it varies for different plant species as seen from field observations (Bolstad and Gower,

1990; Kira et al., 1969; Monsi and Saeki, 1953). Crown area and leaf area index are calculated based on allocation rules and are depending on the annual biomass increment (Sitch et al., 2003). Population density depends on establishment and mortality processes in LPJmL (Sitch et al., 2003).

### 3.2.2.2 *Phenology*

The daily phenology and green leaf status of a PFT ( $Phen_{PFT}$ ) in LPJmL expresses the fractional cover of green leaves (from 0 = no leaves to 1 = full leaf cover). Thus, it represents the temporal dynamic of the canopy greenness. We explored two phenology models in this study: First, we were trying to optimize model parameters of the original phenology module in LPJmL (LPJmL-OP, Sitch et al., 2003). Secondly, we implemented a new phenology module based on the growing season index (GSI) concept (Jolly et al., 2005), hereinafter called LPJmL-GSI.

LPJmL-OP has three different routines for summergreen (i.e. temperature-driven deciduous), evergreen (no seasonal variation) and rain-green (i.e. water-driven deciduous) PFTs (details in Supplement 7.1.1). Obviously, LPJmL-OP misses important controls on phenology like effects of light in all PFTs or effects of water in summergreen and herbaceous PFTs. Additionally, in herbaceous PFTs the end of the growing season is not controlled by environmental conditions but is defined based on fixed calendar dates.

Because of the obvious limitations of LPJmL-OP, we developed the alternative LPJmL-GSI phenology module. The growing season index (GSI) is an empirical phenology model that multiplies limiting effects of temperature, day length and vapour pressure deficit (VPD) to a common phenology status (Jolly et al., 2005). We modified the GSI concept for the specific use in LPJmL (LPJmL-GSI). We defined the phenology status as a function of cold temperature, short-wave radiation and water availability. Additionally to the original GSI model, we added a heat stress limiting function because it has been suggested that vegetation greenness is limited by temperature-induced heat stress in several ecosystems (Bunn et al., 2007; Verstraeten et al., 2006) and has been demonstrated that heat stress reduces plant productivity also without additional water stress (Jiang and Huang, 2001; Van Peer et al., 2004; Poirier et al., 2012). Thus, the daily phenology status of a PFT is the product of the daily cold



temperature ( $f_{cold,PFT}$ ), light ( $f_{light,PFT}$ ), water ( $f_{water,PFT}$ ) and heat stress ( $f_{heat,PFT}$ ) limiting functions:

$$Phen_{PFT} = f_{cold,PFT} \times f_{light,PFT} \times f_{water,PFT} \times f_{heat,PFT} \quad (3.5)$$

Examples for the four functions are shown in Figure 3.2.

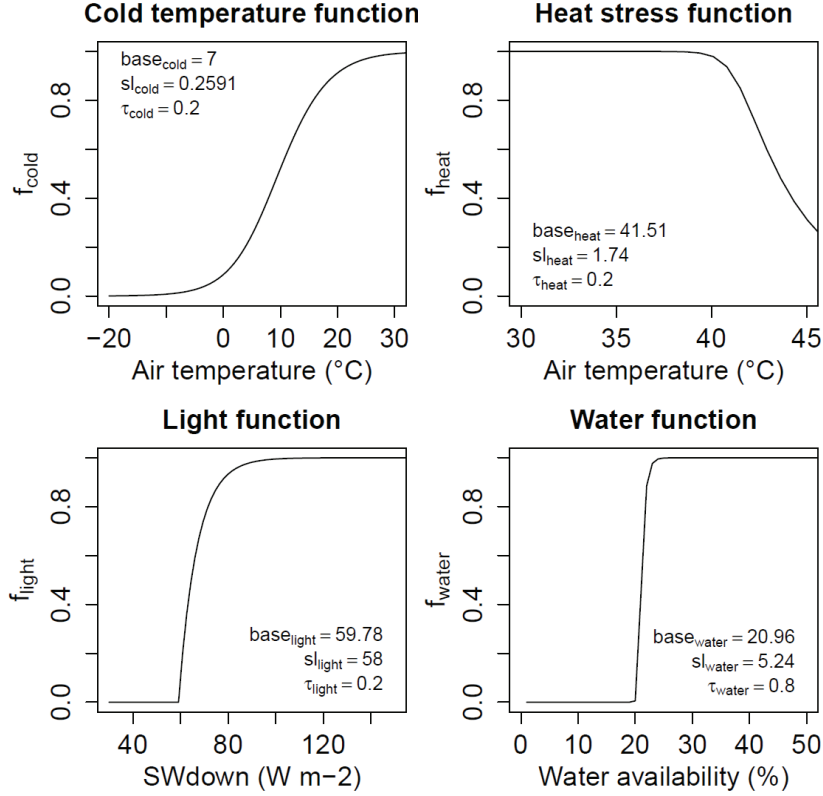


Figure 3.2: Examples of the cold temperature, heat stress, light and water limiting functions for phenology in LPJmL-GSI.

Depending on the chosen parameters the functions have different shapes for each PFT.

The cold temperature limiting function at a daily time step  $t$  is defined as:

$$f_{cold,PFT}^t = f_{cold,PFT}^{t-1} + \left( \frac{1}{1 + e^{-sl_{cold,PFT} \times (T - base_{cold,PFT})}} - f_{cold,PFT}^{t-1} \right) \times \tau_{cold,PFT} \quad (3.6)$$

where  $sl_{cold,PFT}$  and  $base_{cold,PFT}$  are PFT-dependent slope and inflection point parameters of a logistic function based on mean daily air temperature  $T$  (°C). The parameter  $\tau_{cold,PFT}$  is the change rate parameter based on the difference between the actual predicted limiting function value and the previous-day cold temperature limiting function value. This parameter introduces a temporal autocorrelation in the phenology status and avoids abrupt phenological changes because of changing weather conditions.

The light-limiting function was implemented accordingly:

$$f_{light,PFT}^t = f_{light,PFT}^{t-1} + \left( \frac{1}{1 + e^{-sl_{light,PFT} \times (SW - base_{light,PFT})}} - f_{light,PFT}^{t-1} \right) \times \tau_{light,PFT} \quad (3.7)$$

where  $sl_{light,PFT}$  and  $base_{light,PFT}$  are the PFT-dependent slope and inflection point parameters of a logistic function based on daily shortwave downward radiation SW ( $W \text{ m}^{-2}$ ). The parameter  $\tau_{light,PFT}$  is the temporal change rate for the light-limiting function.

The water-limiting function  $f_{water,PFT}$  depends on the daily water availability  $W$  (%) in LPJmL:

$$f_{water,PFT}^t = f_{water,PFT}^{t-1} + \left( \frac{1}{1 + e^{-sl_{water,PFT} \times (W - base_{water,PFT})}} - f_{water,PFT}^{t-1} \right) \times \tau_{water,PFT} \quad (3.8)$$

where  $sl_{water,PFT}$  and  $base_{water,PFT}$  are the PFT-dependent slope and inflection point parameters of a logistic function based on daily water availability.  $W$  is a ratio between water supply from soil moisture and atmospheric water demand (Supplement 7.1.2) (Gerten et al., 2004). The parameter  $\tau_{water,PFT}$  is the temporal change rate for the water-limiting function.

The heat-stress limiting function is defined as the cold-temperature limiting function based on daily air temperature but with a negative slope parameter:

$$f_{heat,PFT}^t = f_{heat,PFT}^{t-1} + \left( \frac{1}{1 + e^{sl_{heat,PFT} \times (T - base_{heat,PFT})}} - f_{heat,PFT}^{t-1} \right) \times \tau_{heat,PFT} \quad (3.9)$$

where  $sl_{heat,PFT}$  and  $base_{heat,PFT}$  are the PFT-dependent slope and inflection point parameters of a logistic function based on  $T$ . The parameter  $\tau_{heat,PFT}$  is the temporal change rate for the heat limiting function.

Besides the additional use of the heat stress limiting function, LPJmL-GSI has important differences to the original GSI phenology model (Jolly et al., 2005). We made the water limiting function dependent on water availability. VPD has been used instead in the original GSI phenology model. Nevertheless, it has been shown that phenology is more driven by soil moisture and plant available water than by atmospheric water demand especially in Mediterranean and grassland ecosystems (Archibald and Scholes, 2007; Kramer et al., 2000; Liu et al., 2013; Yuan et al., 2007) and that GSI performed better when using a soil moisture limiting function instead of the VPD limiting function (Migliavacca et al., 2011). With the implementation of the water limiting function in LPJmL-GSI, phenology depends not only on atmospheric water demand as in the original GSI model but also on water supply from soil moisture.

Additionally, the soil moisture can be modulated through seasonal freezing and thawing in permafrost soils according to the permafrost routines in LPJmL (Schaphoff et al., 2013). Another important difference to the original GSI phenology model is the use of logistic functions instead of stepwise linear functions with fixed thresholds because smooth functions are generally easier to optimize than functions with abrupt thresholds and potentially better represent biological processes. A moving average of 21 days has been used in the original GSI model to create smooth phenological cycles and to avoid abrupt phenology changes because of daily weather variability (Jolly et al., 2005). It has been shown that PFT- and limiting function-dependent time averaging parameters are needed instead of one single time averaging parameter (Stöckli et al., 2011). We implemented change rate parameters  $\tau_{\text{cold}}$ ,  $\tau_{\text{light}}$ ,  $\tau_{\text{water}}$  and  $\tau_{\text{heat}}$  that are PFT- and limiting function-dependent instead of moving average window lengths because LPJmL cannot use the same running window time averaging approach as a prognostic model.

### 3.2.3 Data sets

#### 3.2.3.1 Data sets for parameter optimization: FAPAR, albedo and GPP

We used FAPAR, albedo and GPP data sets to optimize phenology, FAPAR, productivity and vegetation albedo-related parameters in LPJmL (Figure 3.2). We require long-term FAPAR data sets to improve vegetation greenness in LPJmL on seasonal to decadal timescales. Two recently developed data sets provide 30-year time series of FAPAR. The Geoland2 BioPar (GEOV1) FAPAR data set (Baret et al., 2013) (hereinafter called GL2 FAPAR) and the GIMMS3g FAPAR (Zhu et al., 2013) data sets were used in this study.

GL2 FAPAR is defined as the black-sky green canopy FAPAR at 10:15 solar time and has been produced based on SPOT VGT (1999–2012) and AVHRR (1981–2000) observations (Baret et al., 2013). The GL2 FAPAR data set has a temporal resolution of 10 days and a spatial resolution of  $0.05^\circ$  for the AVHRR-period and of  $1/112^\circ$  for the SPOT VGT period. GIMMS3g FAPAR corresponds to black-sky FAPAR at 10:35 solar time and has been produced based on the GIMMS3g NDVI data set (Pinzon and Tucker, 2014; Zhu et al., 2013). GIMMS3g FAPAR has a 15-day temporal resolution and a  $1/12^\circ$  spatial resolution and covers July 1981 to December 2011. We excluded in both

FAPAR data sets observations that were flagged as contaminated by snow, aerosols or clouds. Additionally, we excluded FAPAR observations for months with temperatures  $<0^{\circ}\text{C}$  to exclude potential remaining distortions of snow cover. Both data sets were aggregated to a  $0.5^{\circ}$  spatial and monthly temporal resolution to be comparable with LPJmL simulations. We found that the GL2 AVHRR and GL2 VGT FAPAR data sets have not been well harmonized (Supplement 7.2.1). Thus, we did not use the combined GL2 VGT and AVHRR FAPAR data set for parameter optimization and for analyses of interannual variability and trends but only for analyses and evaluations of mean seasonal cycles and spatial patterns of FAPAR. The GIMMS3g FAPAR data set has no uncertainty estimates. Uncertainty estimates are necessary in multiple data stream parameter optimization to weight single data streams in the total cost function. As a workaround we estimated the uncertainty based on monthly varying quantile regressions to the 0.95 quantile between FAPAR and the FAPAR uncertainty in the GL2 VGT data set. We applied the fitted regressions to the GIMMS3g data set to estimate FAPAR uncertainties (Supplement 7.2.2). The fit to the upper quantile provides conservative uncertainty estimates for the GIMMS3g FAPAR data set.

We used monthly shortwave white-sky albedo time series ranging from 2000–2010 from the MODIS C5 data set (Lucht et al., 2000; Schaaf et al., 2002) to constrain vegetation albedo parameters. Albedo observations in months with  $<5^{\circ}\text{C}$  air temperature and above an albedo of 0.3 were excluded from optimization because we are optimizing only vegetation-related albedo parameters. High albedo values at low temperatures are probably affected by changing snow regimes which is not within our focus of model development and optimization. Thus we are only optimizing growing season albedo. We used mean annual total GPP patterns from the data-oriented MTE (model tree ensemble) GPP estimate (Jung et al., 2011). This GPP estimate uses FLUXNET eddy covariance observations together with satellite observations and climate data to upscale GPP using a machine learning approach (Jung et al., 2011). This data set is not an observation but a result of an empirical model. Nevertheless, evaluation and cross-validation analyses have shown that this data set well represents the mean annual spatial patterns and mean seasonal cycles of GPP whereas it has a poor performance in representing temporal GPP anomalies (trends and extremes) (Jung et al., 2011). Thus, we are only using the mean annual total GPP from this data

set for parameter optimization to constrain LPJmL within small biases of mean annual GPP. We used the mean seasonal cycle from the MTE GPP product as an independent benchmark for model evaluation.

### 3.2.3.2 *Data sets for the prescription of land cover, tree cover and burnt area*

The FAPAR, albedo and GPP data sets do not presumably contain enough information to constrain all processes that control FAPAR dynamics – especially processes like establishment, mortality, competition between PFTs, allocation and disturbances control FPC and thus FAPAR. The optimization of parameters of these processes against appropriate data streams is not feasible within this study. Thus, we directly prescribed land and tree cover fractions as well as burnt areas from observed data to control for some of these processes.

To prescribe land and tree cover in LPJmL, we combined several data sets to create observation-based maps of FPC (Supplement 7.3.1). Land cover maps from remote sensing products are not directly comparable with PFTs in global vegetation models due to differences in classification systems (Jung et al., 2006; Poulter et al., 2011a). PFTs in LPJmL are defined according to biome (tropical, temperate or boreal), leaf type (needle-leaved, broadleaved) and phenology type (summergreen, evergreen, rain green). We extracted the biome information from the Köppen–Geiger climate classification (Kottek et al., 2006) whereas leaf type and phenology were extracted from the SYNMAP land cover map (Jung et al., 2006). FPC was derived from MODIS tree cover (Townshend et al., 2011). Because LPJmL so far classified herbaceous vegetation according to their photosynthetic pathway (i.e. C3, temperate herbaceous and C4, tropical herbaceous), we further subdivided herbaceous PFTs according to biome and introduced a polar herbaceous PFT (PoH) based on the existing temperate herbaceous PFT (TeH) to differentiate tundra from temperate grasslands.

Burnt area data were prescribed directly in LPJmL by combining three data sets, the Global Fire Emissions Database (GFED) burnt area data set (Giglio et al., 2010), the Alaska Large Fire Database (ALFDB) (Frames, 2012; Kasischke et al., 2002) and the Canadian National Fire Database (CNFDB) (CFS, 2010; Stocks et al., 2002). GFED provides monthly burnt area estimates in 0.5° resolution from 1996-2011. Burnt areas from the Alaska (ALFDB) and Canada (CNFDB) fire databases were used to extend

burnt area time series before 1996 for boreal North America. Fire perimeter observations from 1979–1996 from ALFDB and CNFDB were aggregated to  $0.5^{\circ} \times 0.5^{\circ}$  gridded monthly burnt area time series. Observations before 1979 were excluded because fires were not reported for all provinces in Canada. Although the CNFDB contains only fire perimeters  $>200\text{ha}$ , in both databases some fires are missing due to different mapping techniques, and fire perimeters do not agree with burned area, the integration of these data sets provides unique information about spatial-temporal patterns of disturbances especially in boreal ecosystems. It is necessary to simulate fire activity also during the model spin-up as fire influences the equilibrium between vegetation, soil and climate as well. Otherwise biomass would be overestimated at the beginning of the transient model run. For this purpose, we created artificial burnt area time series for the periods 1901–1978 (North America) and 1901–1995 (rest of the world). For this, observed annual total burnt areas from the periods 1979–2011 (North America) and 1996–2011 (rest of the world) were resampled according to temperature and precipitation conditions and assigned to the pre-data period in order to include fire regimes that agree with observed fire regimes in the spinup of LPJmL. This approach assumes that fire regimes in the pre-data period were not different than in the observation period.

### 3.2.3.3 *Data sets for model evaluation*

LPJmL was evaluated against data sets that are independent of the optimization and prescription data sets and against independent temporal or spatial scales of the optimization and prescription data sets. We compared LPJmL against mean annual patterns and mean seasonal cycles of ET from the MTE estimate (Jung et al., 2011). Further, we evaluated model results against spatial patterns of biomass. Ecosystem biomass estimates were taken from satellite-derived forest biomass maps for the tropics (Saatchi et al., 2011) and for the temperate and boreal forests (Thurner et al., 2014) including an estimation of herbaceous biomass (Carvalhais et al., 2014). Additionally, we evaluated LPJmL against independent temporal and spatial scales of the integration data (mean seasonal cycle of GPP, tree cover, inter-annual variability and trends of FAPAR). We were using tree cover from MODIS to evaluate LPJmL model runs with dynamic vegetation.

#### 3.2.3.4 *Climate forcing data and model spin-up*

LPJmL was driven by observed monthly temperature and precipitation data from the CRU TS3.1 data set ranging from 1901–2011 (Harris et al., 2014) as well as by monthly shortwave downward radiation and long-wave net radiation reanalysis data from ERA-Interim (Dee et al., 2011). LPJmL needs a model spin-up to establish PFTs and to bring vegetation and soil carbon pools into equilibrium. The spin-up was performed according to the standard LPJmL modelling protocol (Schaphoff et al., 2013; Thonicke et al., 2010): LPJmL was run for 5000 years by repeating the climate data from 1900–1930. After the spin-up model run, the transient model run was restarted from the spin-up conditions in 1901 and LPJmL was run for the period 1901–2011. Model results were analysed for the observation period (1982–2011). For model optimization experiments we used a different spin-up scheme because the spin-up is computation time demanding and many model runs are needed during optimization experiments. As in the standard modelling protocol, we firstly spin-up the model for 5000 years by repeating the climate from 1901–1930. Secondly, a transient model run was restarted from the spin-up conditions in 1901 and was performed for the period 1901–1979. Thirdly, each optimization experiment was restarted from the conditions in 1979 and a second spin-up for 100 years by recycling the climate from 1979–1988 was performed. The transient model run was restarted from the conditions of the second spin-up and simulated for the period 1979–2011. This second spin-up is needed to bring the vegetation into a new equilibrium which can be caused by a new parameter combination during optimization. From visual analyses of model results, we found that a spin-up time of 100 years for the second spin-up was enough to eliminate trends in FAPAR and GPP that resulted from other equilibrium conditions.

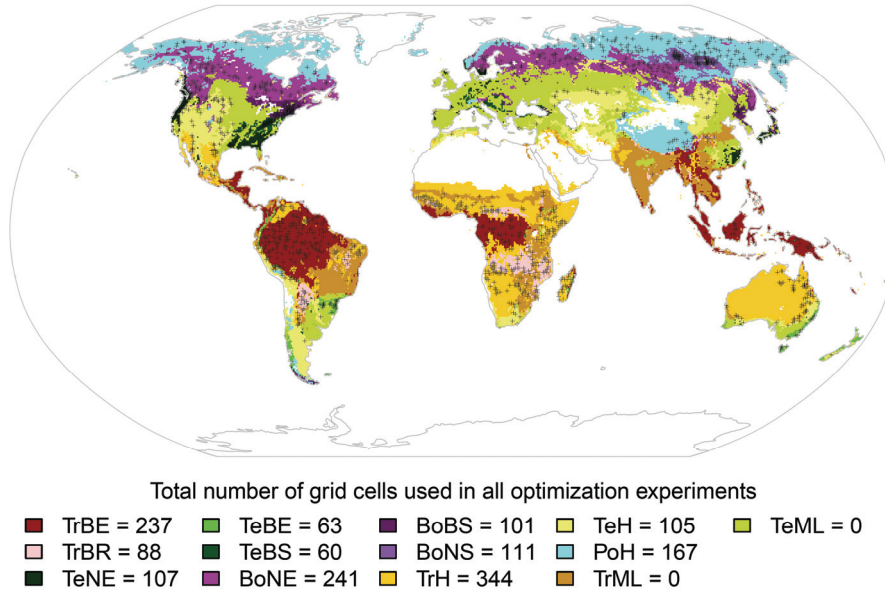


Figure 3.3: Map of the dominant PFT in each grid cell as derived from SYNMAP, Köppen–Geiger climate zones and MODIS VCF.

Grid cells that were used in any of the optimization experiments are shown as black crosses. Some grid cells were used in multiple optimization experiments. Grid cells that are dominated by agriculture were not used for optimization (TrML, tropical managed lands and TeML, temperate managed lands).

### 3.2.4 Model–data integration

#### 3.2.4.1 Prescription of land and tree cover

Land cover is expressed as FPC in LPJmL. We used the observation-based FPC data set to prescribe land and tree cover in LPJmL (Sect. 3.2.3.2, Supplement 7.3.1). The presence of a PFT in a grid cell depends on establishment and mortality in LPJmL (Sitch et al., 2003). A PFT establishes in a grid cell if the climate is within the bioclimatic limits of the PFT for establishment and survival. On the other hand, a PFT dies in a grid cell if the climate is no longer suitable for the PFT. Additionally, mortality occurs because of heat stress, low productivity, competition among PFTs for light, and because of fire disturbance (Sitch et al., 2003; Thonicke et al., 2010).

FPC is the major variable that contributes to inter-annual variability of FAPAR in LPJmL despite the daily phenological status. Thus fixing FPC to the observed value is not a desired solution to prescribe land cover in LPJmL. Fixing FPC would neglect mortality effects on land cover but would also permit the simulation of post-fire succession trajectories. Consequently, we prescribed land cover in LPJmL using a hybrid diagnostic-dynamic approach. In this approach we prescribed the annual



maximum FPC in LPJmL similar to previous approaches (Poulter et al., 2011b). Firstly, we switched off the effects of bioclimatic limits on establishment and mortality. Only these PFTs were allowed to establish in a grid cell that occurred in the observed land cover data set. Vegetation growth depends on the annual biomass increment and allocation rules in LPJmL. This leads to an extension of FPC of each PFT. We limited a further expansion of FPC if the simulated FPC exceeded the observed FPC by replacing the simulated FPC with the observed FPC (prescribed maximum FPC). Consequently, the simulated FPC can be lower than the observed FPC because the PFT is still growing or because the FPC was reduced due to fire, heat stress or low productivity. For herbaceous PFTs we only reduced the FPC if the observed total fractional vegetation cover in a grid cell was exceeded. This allowed herbaceous PFTs to replace tree PFTs if the FPC of trees was reduced due to fire or other mortality effects in the model. With this approach a prescription of land cover can be achieved in LPJmL which well represents observed PFT distributions (Supplement 7.3.2) but still allows for main processes of dynamic vegetation.

#### 3.2.4.2 *Parameter optimization*

Photosynthesis, albedo, FAPAR and phenology-related model parameters of LPJmL were optimized against observed FAPAR and albedo satellite observations and data-oriented estimates of GPP. A description of all parameters including parameter values is given in Supplement 7.4.1. The parameter  $\alpha_a$  is the most important parameter in LPJmL for photosynthesis (Zaehle et al., 2005). This parameter accounts for the amount of radiation that is absorbed at leaf level in comparison to the total canopy. Thus, this parameter is a replacement for a more enhanced model formulation for canopy structure and leaf clumping. We used this parameter to adjust biases in GPP. The PFT-dependent leaf, stem and litter albedo parameters ( $\beta_{\text{leaf}}$ ,  $\beta_{\text{stem}}$  and  $\beta_{\text{litter}}$ ) are mostly sensitive for model simulations of albedo. The parameter  $\beta_{\text{leaf}}$  affects additionally the maximum FAPAR of a PFT. The light extinction coefficient  $k$  controls the FPC of a PFT and thus affects mainly land cover, maximum FAPAR and the available radiation for photosynthesis. All other parameters that were considered in optimization experiments are the parameters of the LPJmL-OP and LPJmL-GSI phenology modules. These parameters contribute mainly to seasonal variations in FAPAR. Some parameters were

excluded from optimization experiments that were identified as insensitive to GPP and FAPAR simulations in PFTs. The temporal change rate parameters cold, light, heat and water are insensitive in most PFTs because of the monthly temporal resolution of the climate forcing data used. The optimization of model parameters was performed by minimizing a cost function between model simulations and observations using a combined genetic and gradient-based optimization algorithm (GENOUD, genetic optimization using derivatives, Mebane and Sekhon, (2011), see Supplement 7.4.2 for details). The cost function  $J$  of LPJmL for a single model grid cell (gc) depends on the scaled model parameter vector  $d$  ( $d = \text{proposed parameter value}/\text{prior parameter value}$ ) and is the sum of square error (SSE) between model simulation and observation weighted by the number of observations (nobs) for each data stream (DS):

$$J(d)_{gc} = \sum_{DS=1}^{DS=n} \frac{SSE_{DS}(d)}{nobs_{DS}} \quad (3.10)$$

The SSE for a single data stream is calculated from the LPJmL simulation of this data stream ( $x_{LPJmL}$ ) and the corresponding observed values ( $x_{obs}$ ) weighted by the uncertainty of the observations ( $x_{obsunc}$ ) for each time step  $t$ :

$$SSE(d) = \sum_{t=1}^{t=n} \frac{(x_{LPJmL,t}(d \times p_0) - x_{obs,t})^2}{x_{obsunc,t}^2} \quad (3.11)$$

where  $p_0$  are LPJmL prior parameters. That means that the minimization of the cost function  $J$  is based on scalars of LPJmL parameters relative to the prior parameter values. Different model optimization experiments were performed for individual grid cell and for multiple grid cells of the same PFT for LPJmL-OP as well as for LPJmL-GSI (Table 3.1). In the grid-cell-based optimization experiments model parameters of the established target tree PFT and the established herbaceous PFT were optimized at the same time. The purpose of grid cell level optimization experiments was to explore the variability of parameters within different regions and PFTs. In the PFT level optimization experiments the cost of LPJmL was calculated as the sum of the cost for each grid cell weighted by the grid cell area  $A$ :

$$J(d)_{PFT} = \frac{\sum_{gc=1}^{gc=n} J(d)_{gc} \times A_{gc}}{\sum_{gc=1}^n A_{gc}} \quad (3.12)$$

For PFT level optimizations parameters of herbaceous PFTs were first optimized for grid cells where only the herbaceous PFT was dominant. In a second step, the optimized parameters of the herbaceous PFTs were used in the optimization of the target tree PFT (Figure 7.9). The purpose of PFT level optimization experiments is to derive optimized parameter sets that can be used for one PFT in global model runs. For grid cell as well as PFT level optimization experiments, we only used grid cells that are vegetated, dominated by one PFT and that are only marginally affected by agricultural use or fire disturbances. These grid cells are called candidate grid cells in the following. We randomly selected grid cells from the set of candidate grid cells to perform grid cell- or PFT level optimization experiments. Table 3.1 gives an overview of all optimization experiments for LPJmL-OP and LPJmL-GSI with the number of used grid cells. Grid cells that were selected for optimization experiments are also shown in Figure 3.3. The PFT level optimization of LPJmL-OP (OP.pft) did not result in plausible posterior parameter sets because of structural limitations of the LPJmL-OP phenology model for herbaceous PFTs (i.e. no water effects, calendar day as end of growing season), raingreen PFT (i.e. binary phenology) and evergreen PFTs (i.e. constant phenology) and was therefore excluded from further analysis. Posterior parameter sensitivities, uncertainties and correlations were explored by analysing the maximum likelihood and the posterior range of each parameter as derived from all parameter sets from the genetic optimization algorithm (Supplement 7.4.3).

#### 3.2.4.3 *Model evaluation and time series analysis*

Global model runs of LPJmL were performed in order to evaluate model results against the integration data, against independent metrics of the integration data and against independent data streams. We evaluated results from LPJmL-OP with standard parameters (LPJmL-OP-prior), from LPJmL-OP with optimized productivity, albedo and FAPAR parameters from grid cell level optimization experiments (LPJmL-OP-gc) and from LPJmL-GSI with optimized parameters from PFT level optimization

experiments (Table 3.2). We did not use optimized phenology parameters in the LPJmL-OP-gc model run because we were not able to derive plausible phenology parameters in optimization experiments of LPJmL-OP. All model runs were performed with dynamic vegetation and prescribed burnt areas.

Table 3.1: Overview of optimization experiments with information sources for prior and posterior parameter sets. Parameter values and prior parameter ranges for each parameter set are listed in Supplement 7.4.1.

Experiment	Description	Number of randomly selected grid cells	Prior parameter set and sources	Posterior parameter set
OP.prior	Parameters or model results of LPJmL-OP with standard parameters	--	Table 7.2 Sitch et al. (2003): $\alpha_a$ , $k$ , $ramp$ , $aphen_{min}$ , $aphen_{max}$ , $Wscal_{min}$ Strengers et al. (2010): $sfc$ and albedo parameters (partly estimated from MODIS albedo)	--
OP.gc	Optimization of single grid cells of LPJmL-OP.	530 in total TrBE 66, TrBR 51, TeNE 46, TeBE 32, TeBS 32, BoNE 68, BoBS 40, BoNS 49, TeH 66, TrH 80	Table 7.2 Parameters as in OP.prior	One optimized parameter set per grid cell. Median-averaged values for PFTs (Table 7.3)
OP.pft (results not shown)	Optimization of multiple grid of LPJmL-OP. Multiple grid cells of the same dominant PFT were optimized at the same time.	673 in total TrBE 50, TrBR 80, TeNE 50, TeBE 50, TeBS 80, BoNE 50, BoBS 80, BoNS 158, TeH 50, TrH 25	Median-averaged values for PFTs from posterior values of OP.gc (Table 7.3)	-- (No useful posterior parameter sets were found)
GSI.prior	Parameters or model results of LPJmL-GSI with standard parameters.	--	Table 7.4 OP.gc: $\alpha_a$ , $k$ , $sfc$ , $\beta_{leaf}$ , $\beta_{litter}$ , and $\beta_{stem}$ Stöckli et al. (2011): parameters for cold and light limiting functions derived from fitting logistic functions to stepwise functions as reported in Stöckli et al. (2011)	--
GSI.gc	Optimization of single grid cells of LPJmL-GSI.	348 in total TrBE 33, TrBR 33, TeNE 32, TeBE 22, TeBS 43, BoNE 30, BoBS 41, BoNS 30, TeH 46, TrH 38	Parameters as in GSI.prior (Table 7.4)	One optimized parameter set per grid cell.
GSI.pft	Optimization of multiple grid of LPJmL-GSI. Multiple grid cells of the same dominant PFT were optimized at the same time.	500 in total TrBE 30, TrBR 30, TeNE 30, TeBE 30, TeBS 30, BoNE 50, BoBS 30, BoNS 60, TeH 70, TrH 70, PoH 70	Parameters as in GSI.prior (Table 7.4)	Table 7.5 (one optimized parameter set per PFT)

We aggregated monthly FAPAR time series to mean annual FAPAR to evaluate inter-annual variability and trends. Mean annual FAPAR time series were averaged from all monthly values with mean monthly air temperatures  $>0^\circ\text{C}$  to exclude potential remaining effects of snow in the observed FAPAR time series. Trends in mean annual

FAPAR time series and trend breakpoints were computed using the “greenbrown” package for the R software (Forkel et al., 2013). In this implementation, trends are computed by fitting piece-wise linear trends to the annual FAPAR time series using ordinary least squares regression. The significance of trends was computed using the Mann–Kendall trend test (Kendall, 1975; Mann, 1945).

Table 3.2: Overview of global model runs that were used in this study for model evaluation.

Model run	Phenology model	Parameter set	Further settings
LPJmL-OP-prior	original phenology	LPJmL standard parameters as in the OP.prior experiment (Table 7.2)	dynamic vegetation/no prescribed land cover, prescribed agricultural land use, prescribed observed burnt area
LPJmL-OP-gc	original phenology	Optimized productivity, FAPAR and albedo parameters from the OP.gc optimization experiment, but original phenology parameters as in the OP.prior experiment (Table 7.3)	
LPJmL-GSI	GSI-based phenology	Parameters from the GSI.pft optimization experiment (Table 7.5)	

### 3.3 Results and discussion

#### 3.3.1 Parameter optimization

##### 3.3.1.1 Performance of phenology models

The newly developed LPJmL-GSI phenology model resulted in significantly higher correlations with monthly GIMMS3g FAPAR than LPJmL-OP in all PFTs except in the tropical broadleaved evergreen (TrBE) and boreal broadleaved summergreen (BoBS) PFTs (Figure 3.4). LPJmL-OP with prior parameters had high correlations with monthly GIMMS3g FAPAR in broadleaved summergreen PFTs (TeBS median  $r = 0.87$ , BoBS median  $r = 0.92$ ) and medium correlations in boreal needle-leaved PFTs (BoNE median  $r = 0.53$ , BoNS median  $r = 0.6$ ). In all other PFTs, LPJmL-OP had low correlations with monthly GIMMS3g FAPAR. The correlation against monthly GIMMS3g FAPAR did not significantly improve in all PFTs after grid cell level optimization experiments of LPJmL-OP (Figure 3.4). The use of the newly developed LPJmL-GSI phenology model already significantly improved the correlation with monthly GIMMS3g FAPAR in all PFTs except in the temperate herbaceous (TeH) and BoBS PFTs. LPJmL-GSI had significantly higher correlations with monthly GIMMS3g FAPAR after grid cell level optimization experiments in the TrBR, TeNE, TeBS, TeH, BoBS and BoNS PFTs. After PFT level optimization experiments, LPJmL-GSI had median correlation coefficients  $> 0.5$  in all PFTs except in broadleaved evergreen PFTs

(TrBE, TeBE). These results prove that the raingreen, evergreen and herbaceous phenology schemes of LPJmL-OP were not able to reproduce temporal FAPAR dynamics despite the attempt of parameter optimization and that LPJmL-GSI can reproduce seasonal FAPAR dynamics in most PFTs.

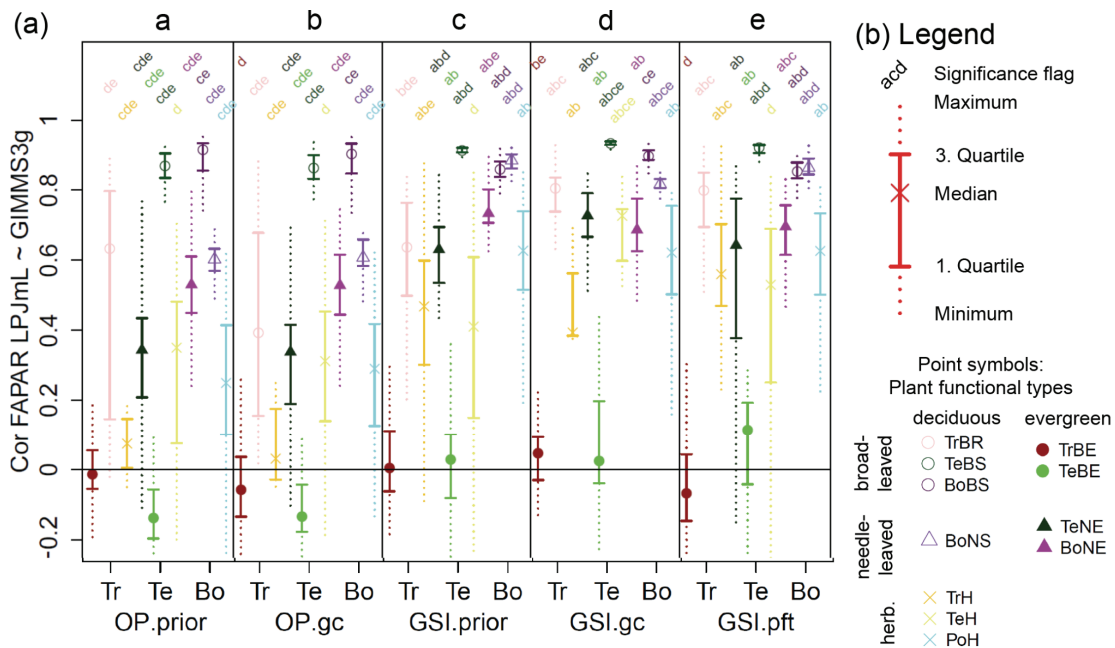


Figure 3.4: Distribution of the correlation coefficient between monthly LPJmL and GIMMS3g FAPAR (1982–2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes.

Correlation coefficient for LPJmL-OP with default parameters (a, OP.prior), after grid cell level optimizations (b, OP.gc); cost for LPJmL-GSI with prior parameters (c, GSI.prior), after grid cell level optimizations (d, GSI.gc) and after PFT level optimizations (e, GSI.pft). Biomes are Tr (tropical), Te (temperate) and Bo (boreal/polar). Each distribution is plotted according to usual boxplot statistics. The point symbols indicate the plant functional type. The significance flag on top of each distribution shows if a distribution is significant different ( $p \leq 0.01$ ) to the corresponding distribution of the same PFT in another optimization experiment. The significance is based on the Wilcoxon rank-sum test. For example “acd” indicates a significant difference to the main categories (a) (OP.prior), (c) (GSI.prior) and (d) (GSI.gc) but no significant difference to (b) (OP.gc) and (e) (GSI.pft).

The low correlation coefficients between LPJmL-GSI and GIMMS3g FAPAR after optimization experiments in broadleaved evergreen PFTs (TrBE, TeBE) might be caused by the specific properties of the FAPAR data set in these PFTs. GIMMS3g FAPAR does not have a clear seasonal cycle but a high short-term variability in broadleaved evergreen forests. These regions are often covered by clouds that inhibit continuous optical satellite observations. The high short-term variability results ultimately in low correlation coefficients between both LPJmL versions (LPJmL-OP and

LPJmL-GSI) and GIMMS3g FAPAR time series. In temperate broadleaved evergreen forests, the GIMMS3g FAPAR data set might have a wrong seasonality. In these regions, the mean seasonal FAPAR cycles from the GIMMS3g and GL2 VGT FAPAR data sets are anti-correlated and FAPAR from LPJmL-GSI agrees better with the GL2 VGT data set. Because of these reasons, we did not expect to improve seasonal FAPAR dynamics in broadleaved evergreen forests with the current model–data integration setup.

All optimization experiments of LPJmL-OP and LPJmL-GSI resulted in a significant reduction of the cost in comparison to the respective prior models (Supplement 7.4.4, Figure 7.10). Nevertheless, the prior parameter set of LPJmL-GSI resulted already in a significant lower cost than the grid cell level optimized parameter sets of LPJmL-OP in tropical and polar herbaceous PFTs, and in temperate broadleaved summergreen and boreal needle-leaved summergreen PFTs. The reduction of the overall cost was in all model optimization experiments usually associated with a significant reduction of the annual GPP bias (Figure 7.11). LPJmL-OP with prior parameters underestimated mean annual GPP in the tropical broadleaved evergreen PFT and overestimated mean annual GPP in all other PFTs. Grid cell level optimization experiments of LPJmL-OP resulted in a significant reduction of the GPP bias in all PFTs except in the polar herbaceous PFT (PoH). We were not able to remove the GPP bias and to reduce the cost of LPJmL-OP and of LPJmL-GSI in the PoH PFT in optimization experiments because of inconsistencies between the FAPAR and GPP data sets or in the LPJmL formulation. LPJmL was not able to sustain the relatively high peak FAPAR in tundra regions as seen in the GIMMS3g data set given the low mean annual GPP of the MTE data set (Supplement 7.4.4). These inconsistencies might be related to higher uncertainties of the GPP and FAPAR data sets in tundra regions where the MTE GPP data set is not covered by many eddy covariance measurement sites, and where satellite-based FAPAR observations are affected from high sun zenith angles (Tao et al., 2009; Walter - Shea et al., 1998). On the other hand, dominant tundra plant communities like mosses and lichen are not represented in LPJmL (Supplement 7.4.4). All model optimizations experiments kept growing season albedo within reasonable ranges in comparison to MODIS albedo (Figure 7.12). These results demonstrate an improved performance of optimized model parameter sets over prior model parameter sets and

of LPJmL-GSI over LPJmL-OP regarding a cost that is defined based on 30 years of monthly FAPAR, mean annual GPP and 10 years of monthly vegetation albedo.

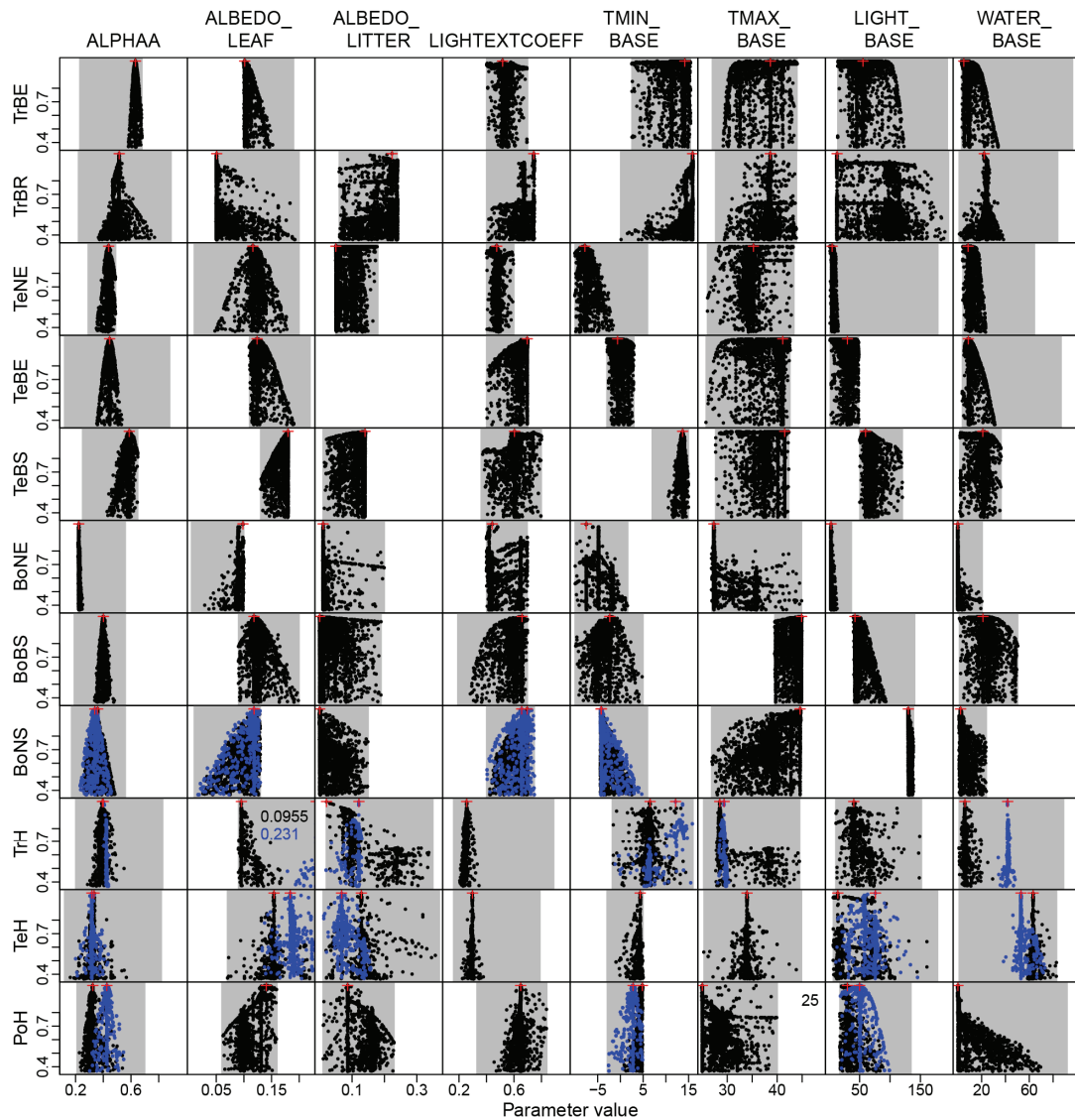


Figure 3.5: Uncertainty and sensitivity of LPJmL-GSI parameters derived from all individuals of genetic optimizations at PFT level.

Shown is the relationship between parameter values and the likelihood of the corresponding parameter vector. The likelihood is normalized with the likelihood of the optimum parameter set. Only individuals with  $dAIC < 2$  are shown. Grey areas indicate the uniform prior parameter range. Red crosses indicate the optimum parameter value. The optimum parameter value is indicated as text in a plot if it is outside of the plotting range. Results from two independent optimization experiments are shown for the BoNS, TrH, TeH and PoH PFTs (black and blue colours, respectively) but not all parameters were included in both experiments. The parameter ALBEDO\_LITTER in the TrBE and TeBE PFTs was not considered in optimization experiments.



### 3.3.1.2 *Parameter sensitivities and uncertainties*

The uncertainty of productivity and albedo-related parameters was reduced after optimization of LPJmL-GSI in most PFTs while the reduction of the uncertainty of phenology-related parameters depended often on plant functional type (Figure 3.5). Prior and posterior parameter values from each optimization experiment are listed in the Supplement (Tables 7.2 to 7.5).

The parameter  $\alpha_a$  (absorption of light at leaf level in relation to canopy level) was sensitive within a narrow parameter range for all PFTs. The posterior  $\alpha_a$  parameter range was smaller than the uniform prior range in all PFTs. In all optimization experiments we found for the parameter  $\alpha_a$  a gradient from high values in tropical to low values in boreal PFTs (Figure 7.13). This pattern reflects the initial overestimation of mean annual GPP in temperate and boreal PFTs and underestimation of GPP in tropical regions with the prior parameter set of LPJmL-OP. Thus, the low  $\alpha_a$  parameter values probably account for nitrogen limitation effects on productivity in boreal forests (Vitousek and Howarth, 1991) that are currently not considered in LPJmL. A future implementation of nitrogen limitation processes in LPJmL requires a re-optimization of the  $\alpha_a$  parameter.

The leaf albedo parameter  $\beta_{\text{leaf}}$  was sensitive in all PFTs and the posterior  $\beta_{\text{leaf}}$  parameter range was smaller than the prior parameter range in evergreen PFTs. In these evergreen PFTs the  $\beta_{\text{leaf}}$  parameter was well constrained because albedo satellite observations are less affected by variations in background albedo (soil, snow) than in deciduous PFTs. In all other PFTs the  $\beta_{\text{leaf}}$  posterior parameter range was equal the prior parameter range or the optimized parameter value was close to a boundary of the prior parameter range. This result indicates that the albedo routines in LPJmL should consider variations in background albedo caused by changes in soil properties, soil moisture, or snow conditions in order to accurately reproduce satellite-observed albedo time series (see discussion in Supplement 7.4.5). Nevertheless, the optimization of the leaf albedo parameter  $\beta_{\text{leaf}}$  resulted in values that differed especially between broadleaved and needleleaved evergreen PFTs as well as herbaceous PFTs (Figure 3.5, Figure 7.14). Low leaf albedo parameters in needle-leaved evergreen PFTs (TeNE and BoNE) and high leaf albedo parameters in broadleaved summergreen and herbaceous PFTs agree well with the patterns reported by Cescatti et al. (2012).

The light extinction coefficient  $k$  was sensitive for all PFTs but the posterior parameter range was only in herbaceous PFTs and in the BoBS PFT smaller than the prior parameter range (Figure 3.5). In all PFTs this parameter had a large spatial variability (Figure 7.15). The parameter  $k$  affects mostly the FPC and thus the maximum FAPAR. Thus, this parameter cannot be well constrained for tree PFTs in the current optimization setup because the maximum FPC of trees was prescribed from the land and tree cover data set. On the other hand, the maximum FPC of herbaceous PFTs was not prescribed from observations which resulted in narrow  $k$  posterior parameter ranges for herbaceous PFTs. The parameter  $k$  was optimized towards a very high value in the BoNS PFT ( $k = 0.7$ ) due to high tree mortality rates after low productivity years (Supplement 7.4.5). This parameter would result in an overestimated PFT coverage in model runs with dynamic vegetation. Thus, we performed a second optimization experiment for this PFT (blue in Figure 3.5) where  $k_{\text{BoNS}}$  was limited to 0.65. This optimization experiment resulted in similar posterior values for the other parameters. Although the  $k$  parameter was well constrained for the TrH, TeH and PoH PFTs, these parameters cannot be used in the final parameter set of LPJmL-GSI. In dynamic vegetation model runs, the relatively low  $k$  parameter values for the TrH and TeH PFTs and relatively high values for the PoH PFT would result in an underestimation of herbaceous coverage in temperate and tropical climates and an overestimation of herbaceous coverage in boreal and polar climates, respectively. Therefore, we performed three more optimization experiments for herbaceous PFTs where we fixed  $k$  at 0.5 (blue in Figure 3.5). These optimization experiments resulted in similar  $\alpha_a$  parameters but different albedo parameters and phenology parameters in order to compensate for biases in FAPAR and albedo that were introduced by the fixed  $k$  parameter. Thus, the high spatial variability and the large uncertainty of the light extinction coefficient  $k$  require re-addressing this parameter in a model optimization setup with dynamic vegetation using tree and vegetation cover data or perhaps a replacement by a better representation of canopy architecture and radiative transfer.

The sensitivity and posterior uncertainty of phenology-related model parameters depended often on plant functional type. The parameter  $\text{base}_{\text{cold}}$  which controls the effect of cold temperature on phenology was sensitive in all PFTs except the TrBE and TrH PFTs. The posterior parameter range was smaller than the prior parameter range

in temperate PFTs (TeNE, TeBS and TeH). The parameter  $\text{base}^{\text{heat}}$  which controls the effect of heat stress on phenology was sensitive in TrBR, TrH, TeH, BoNE and BoNS PFTs while in other PFTs this parameter was only sensitive towards the boundaries of the prior parameter range. Nevertheless, the posterior parameter range was only smaller than the prior parameter range in TrBR and TrH PFTs. The parameter  $\text{base}^{\text{light}}$  was sensitive in temperate and boreal PFTs. In tropical PFTs this parameter is only sensitive above a certain threshold (i.e.  $60 \text{ W m}^{-2}$  for TrBE and  $100 \text{ W m}^{-2}$  for TrBR). The parameter  $\text{base}^{\text{water}}$  was sensitive in all PFTs. The posterior parameter range of this parameter was smaller in all PFTs except in TeBS, BoNE, BoBS and BoNS PFTs. Although the parameter  $\text{base}^{\text{water}}$  had a large variability among PFTs, it was generally optimized towards higher values in PFTs that are presumably water controlled (TrBR, TeBS, TrH, TeH) and optimized towards lower values in PFTs that are presumably less water controlled (TrBE, TeNE, BoNE, BoNS, PoH). This result indicates that FAPAR of water-controlled PFTs reacts already to small decreases in water availability whereas other PFTs react only to strong decreases in water availability. We found no strong correlations between posterior values of the phenology-related model parameters (Figure 7.16) which indicates the ability to disentangle the relative effects of temperature, light and water on phenology. As the  $\text{base}^{\text{water}}$  parameter was the only phenology parameter which was sensitive in all PFTs, this indicates that water availability is the only phenological control that acts in all PFTs.

### **3.3.2 Effects of an improved phenology module in LPJmL**

#### *3.3.2.1 Effects on seasonal and inter-annual greenness dynamics*

LPJmL-GSI represents better the observed spatial patterns and seasonal-to-decadal temporal dynamics of vegetation greenness (FAPAR) than LPJmL-OP (Figure 3.6, Supplement 7.5.3). Whereas LPJmL-OP overestimated mean annual FAPAR in many high-latitude and semi-arid regions, LPJmL-GSI was closer to both data sets and within the uncertainty of the GL2 VGT FAPAR data set in most regions and under most climate conditions (Figure 7.22). LPJmL-GSI still overestimated mean annual FAPAR in temperate dry regions, but this overestimation was reduced in comparison to LPJmL-OP. We further observe a substantial improvement in LPJmL-GSI regarding the seasonal cycles, monthly and annual dynamics of FAPAR as retrieved from the

GIMMS3g and GL2 VGT FAPAR data sets (Figure 3.6, Figures A23–A25). Monthly FAPAR time series from LPJmL-GSI were significantly ( $p \leq 0.05$ ) higher correlated with GIMMS3g than from LPJmL-OP in boreal forests of eastern Siberia, in the North American tundra, in temperate and tropical grasslands of central Asia, North America, Australia and especially, in the Sahel (Figure 3.6a). This is because of an improved representation of spring onset and the end of the growing season in temperate and boreal forests and in herbaceous PFTs (Figure 7.24). The highest differences between simulated and observed mean seasonal FAPAR cycles were observed in the temperate broadleaved evergreen PFT, where both model versions had opposite, although insignificant, relationships to the GIMMS3g data sets. For this PFT, the observational constraints are particularly problematic, where a weak agreement and opposite relationship is observed between the two data sets ( $r = -0.48$ ). Globally, LPJmL-GSI describes better the inter-annual dynamics of GIMMS3g FAPAR when compared to the previous model versions (Figure 3.6). In 20% of the land the difference to other model versions is statistically significant, and in 40% does not detract from the previous model versions. This improvement in inter-annual variability is especially seen in temperate and tropical dry regions, with sparse tree cover and grassland dominated ecosystems (western United States, central Asia, the Sahel, southern Africa, and Australia) (Figure 7.25). In the Arctic, boreal and temperate climates LPJmL usually shows a higher correlation with the GIMMS3g data set than the correlation observed between both data sets (GIMMS3g and GL2 VGT). These results demonstrate that LPJmL-GSI can explain the inter-annual variability of the GIMMS3g FAPAR data set especially in temperate and boreal forests and temperate and tropical grasslands. Overall, the global spatial representation of phenological dynamics in LPJmL-GSI improves significantly over the previous model versions from seasonal to inter-annual timescales. Given the inclusion of water controls on phenological development, these results emphasize the importance of water availability in explaining the mean spatial patterns of vegetation greenness, but also the seasonal phenology as well as inter-annual dynamics in vegetation development.

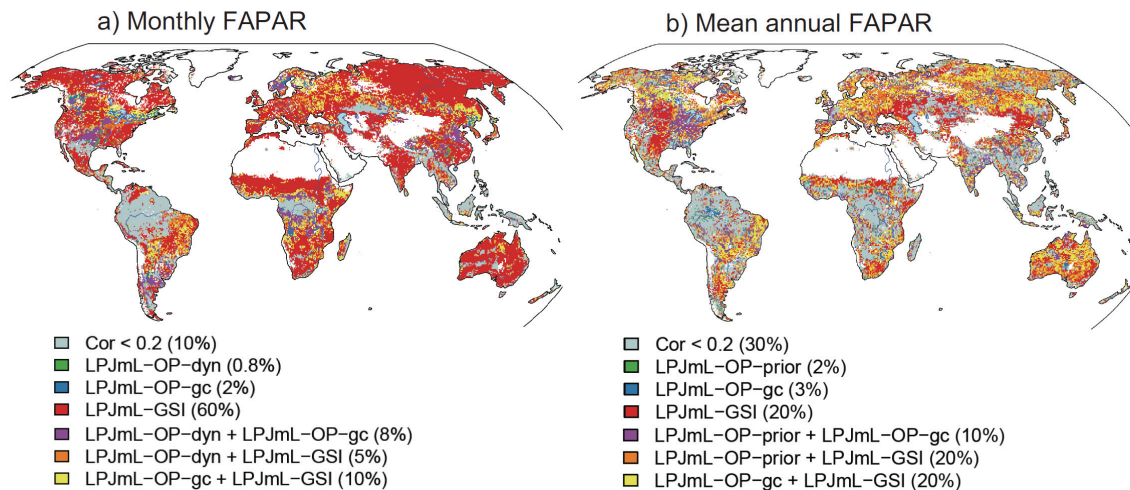


Figure 3.6: Best LPJmL model runs for (a) monthly FAPAR dynamics (1982–2011,  $n = 360$  months) and (b) time series of mean annual FAPAR (1982–2011,  $n = 30$  years).

The best LPJmL model run has the highest correlation coefficient between simulated LPJmL FAPAR and GIMMS3g FAPAR. If one model run is shown the correlation coefficient of this best model is significantly higher than that of the second best model run ( $p \leq 0.05$ , Fisher z transformation on difference in correlation). If two model runs are shown the correlation coefficients of the first and second best model runs are not significantly different from each other ( $p > 0.05$ ).

### 3.3.2.2 *Effects on trends in vegetation greenness*

The role of different climate drivers underlying the greening and browning trends in vegetation activity is still highly debated and the dominant factors show a strong spatial variability (de Jong et al., 2013a). The consideration of different environmental controls on the phenological development in LPJmL shows a significant improvement in representing such dynamics when compared to the previous model formulations (Figure 3.7).

Both LPJmL-OP and LPJmL-GSI reproduced the observed greening trends in tundra regions and in boreal forests of Siberia. In both model versions this greening is mostly driven by annual changes in foliar projective cover and effects of temperature on spring phenology. This agrees with observational studies that identified temperature increases as drivers for an increasing shrub cover in tundra ecosystems (Blok et al., 2011a; Forbes et al., 2010; Myers-Smith et al., 2011; Raynolds et al., 2013; Sturm et al., 2001) and that found a positive association between warming, increasing tree ring widths and NDVI greening in boreal forests of eastern Siberia (Berner et al., 2011, 2013).

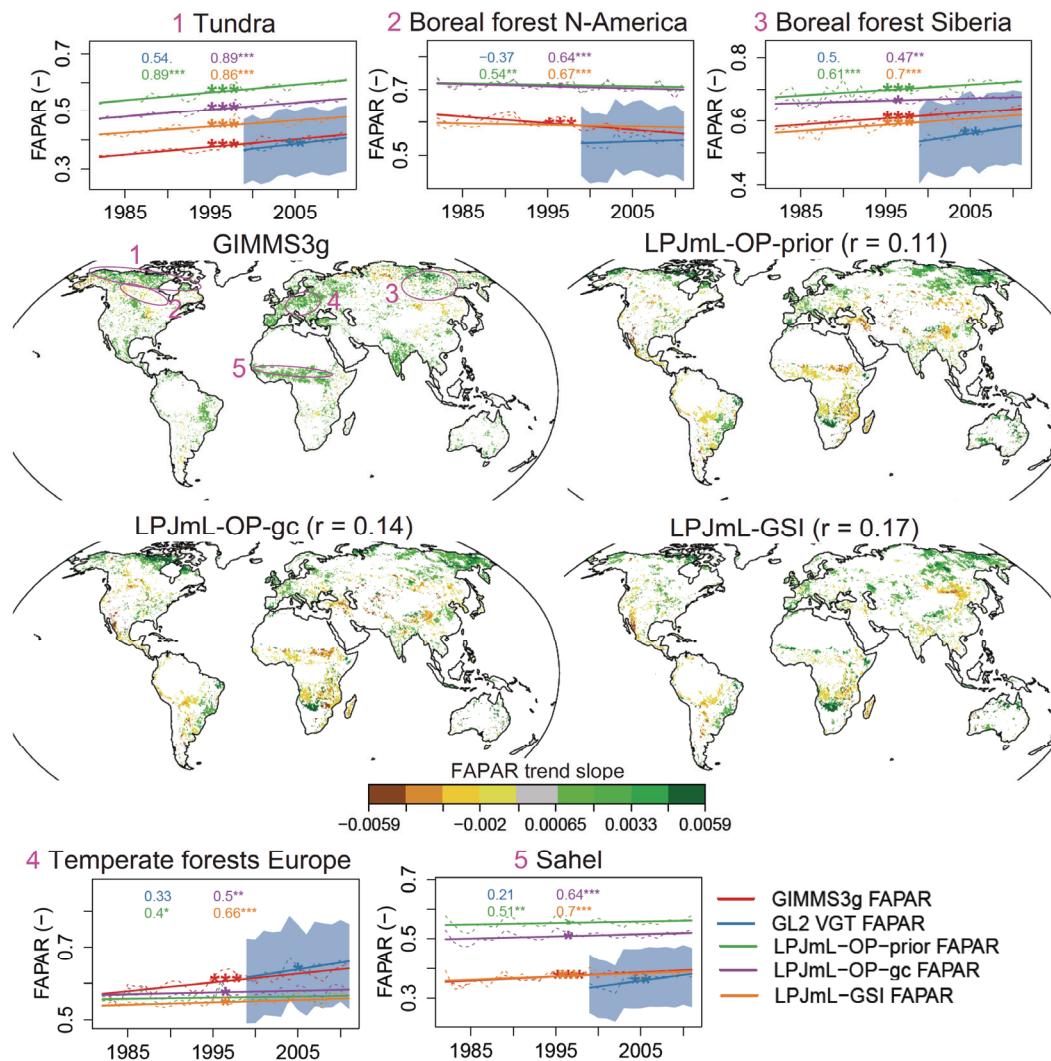


Figure 3.7: Comparison of trends in mean annual FAPAR from LPJmL and from satellite data. Trends were computed between 1982 and 2011 as linear trends. The significance of a trend was determined using the Mann–Kendall trend test. Only significant trends slopes ( $p \leq 0.05$ ) are displayed in each map. Spatial correlations of trend slopes (Spearman coefficient) between LPJmL and the GIMMS3g data set are given in the map titles. Time series are showing mean annual FAPAR time series and trends spatially averaged for the regions as indicated in the first map. The blue area in time series represents the uncertainty of the GL2 VGT FAPAR data set. Numbers in the time series plot are correlation coefficients between mean annual FAPAR time series from GIMMS3g and from GL2 or LPJmL model runs, respectively. The significance of a trend and of the correlation is indicated as point symbol: \*\*\*  $p \leq 0.001$ , \*\*  $p \leq 0.01$ , \*  $p \leq 0.05$ , .  $p \leq 0.1$ .

Parts of the boreal forests in North America showed significant browning trends in the GIMMS3g data set but a tendency to positive trends in the GL2 data set. The simulation results from LPJmL-GSI are in agreement with the GIMMS3g-based browning trends, rather than greening trends. However, these model-based browning trends were not as strong as in the GIMMS3g data set. In LPJmL-GSI these browning trends are caused by the effects of seasonal light and water effects on phenology, and

by fire activity. In the GIMMS3g data set these browning trends were related to several environmental factors like fire activity (Goetz et al., 2005), temperature-induced drought stress (Beck et al., 2011; Bunn and Goetz, 2006) and to snow-regulated changes in soil water availability (Barichivich et al., 2014).

The Sahel had widespread greening trends in the GIMMS3g FAPAR data set. Whereas LPJmL-OP simulated browning trends, the implementation of water availability effects on phenology enabled LPJmL-GSI to reproduce the observed greening trends. Increases in precipitation and rain-use efficiency were also identified in observational studies as the main drivers of positive trends in vegetation greenness in the Sahel (Fensholt et al., 2013).

Overall, we observed that both LPJmL-OP and LPJmL-GSI reproduced the greening trends in tundra, boreal and temperate forests, although LPJmL-GSI showed a wider agreement in the extent of browning trends in the boreal forests of North America. Further, in the Sahel region, the greening trends can only be reproduced through the inclusion of water availability controls on the phenology development. These results demonstrate that environmental controls like light, heat stress and water availability contribute to a better description of regional greening and browning trends in very different bioclimatic regions of the globe. Hence, a comprehensive characterization of the different environmental controls on phenological development is essential in performing model-based analysis of long-term trends in vegetation activity.

### 3.3.2.3 *Effects on carbon fluxes and stocks*

LPJmL-GSI and LPJmL-OP-gc with optimized parameters represented better the global patterns and mean seasonal cycles of gross primary production and biomass than LPJmL with original phenology and prior parameters (LPJmL-OP-prior) (Figure 3.8). LPJmL-OP-prior overestimated mean annual GPP and biomass in most polar, boreal and temperate regions. LPJmL-OP-prior underestimated mean annual GPP but overestimated mean annual biomass in tropical regions around the equator. These biases were reduced in LPJmL-OP-gc and LPJmL-GSI. LPJmL generally overestimated GPP also in arid regions but these biases were reduced after optimization in LPJmL-OP-gc and LPJmL-GSI (Figure 7.18). We also found that the mean seasonal cycle of GPP from LPJmL-GSI agreed better with the mean seasonal GPP cycle from the MTE

estimate especially in temperate forests and in tropical, temperate and polar grasslands (Supplement 7.5.1, Figure 7.17) although no information about the seasonality of GPP was included in optimization experiments. LPJmL-GSI still overestimated biomass in some tropical regions (African Savannas, southeast Brazil, south and south-east Asia) (Figure 7.19). These regions were mainly simulated as agricultural lands in LPJmL, i.e. as different crop functional types (CFTs). The LPJmL-GSI phenology module was not applied or optimized for agricultural regions, where the seasonal phenological development is prescribed according to the CFTs parameterizations from Bondeau et al. (2007). Generally, LPJmL-GSI performed substantially better than LPJmL-OP-prior and LPJmL-OP-gc when comparing the global total carbon fluxes and stocks to the data-oriented estimates (Supplement 7.5.1, Table 7.6). These results demonstrate that in addition to the optimization of productivity parameters in LPJmL, the implementation of the new GSI-based phenology improved estimates of spatial patterns, seasonal dynamics, and global totals of gross primary production and biomass.

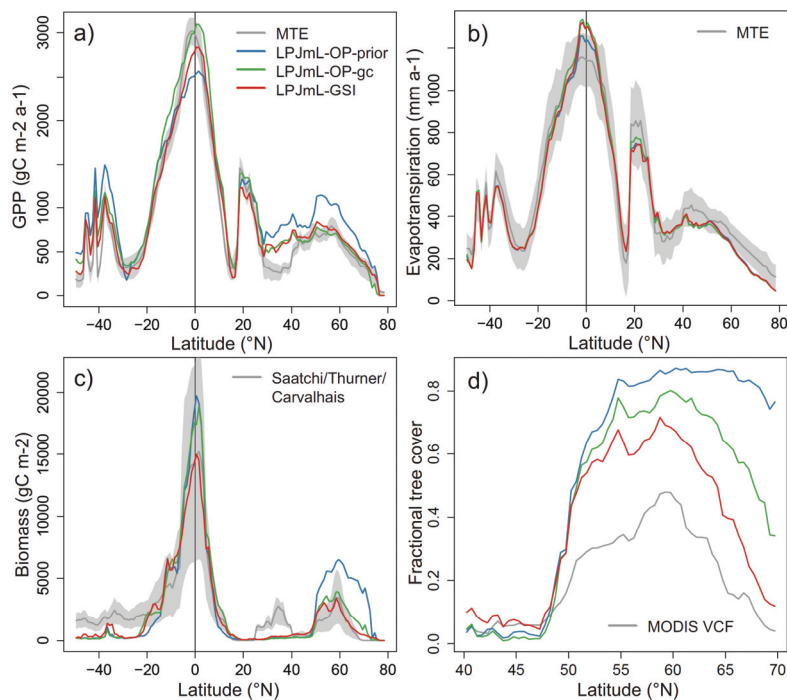


Figure 3.8: Latitudinal gradients of (a) gross primary production (GPP), (b) evapotranspiration, (c) biomass and (d) tree cover from data-oriented estimates and from LPJmL model simulations.

Gradients were spatially averaged (median) from all 0.5° grid cells for latitudinal bands of 1° width. Grey areas represent uncertainty estimates for the data-oriented estimates.



#### 3.3.2.4 *Effects on forest distribution*

LPJmL-GSI with dynamic vegetation better represented the observed tree cover in high-latitude regions than LPJmL-OP-prior and LPJmL-OP-gc (Figure 3.8d). LPJmL-OP-prior highly overestimated tree cover in boreal and Arctic regions and simulated a too northern Arctic tree line in comparison with tree cover from MODIS observations. Although this overestimation was reduced after optimization, LPJmL-OP-gc still highly overestimated tree cover in boreal and temperate regions. The occurrence of trees was shifted southwards in LPJmL-GSI. Although LPJmL-GSI still overestimated tree cover in boreal regions, this overestimation was much lower than in LPJmL-OP-gc. LPJmL-OP-prior and LPJmL-OP-gc slightly underestimated tree cover in temperate regions around 45°N but this was well reproduced by LPJmL-GSI. We found no differences in tree cover between LPJmL-OP and LPJmL-GSI in other parts of the world where tree cover is highly affected from agricultural land use and thus implicitly prescribed to LPJmL. These results demonstrate that additional to the optimization of productivity parameters in LPJmL-OP-gc, the newly developed GSI-based phenology model and the optimized model parameters contribute to a better representation of tree cover in high-latitude regions.

#### 3.3.2.5 *Effects on evapotranspiration processes*

Evapotranspiration from LPJmL agreed well with the data-oriented MTE estimates (Figure 3.8b). The implementation and optimization of the new GSI-based phenology did not affect ET much. ET increased only in tropical rainforests around the equator in LPJmL-GSI and LPJmL-OP-gc in comparison to LPJmL-OP-prior because of the increased GPP in these regions. In other regions ET remained almost unchanged. But this does not imply that the structural improvements in LPJmL-GSI did not affect the transpiration processes (Figures. A20, A21). Indeed, LPJmL-GSI had lower interception losses than LPJmL-OP in boreal forests because of the reduced tree cover. On the other hand this implies that simulated soil evaporation was increased. Furthermore, interception and soil evaporation had slightly shifted seasonal cycles in LPJmL-GSI compared to LPJmL-OP due to the seasonal differences in timing of leaf development and senescence stages (Figure 7.21). Consequently, small differences in total

evapotranspiration result from opposite and compensatory changes in interception and soil evaporation and slight changes in transpiration fluxes in LPJmL-GSI.

### 3.3.3 Applicability and challenges of the LPJmL-GSI phenology module

The LPJmL-GSI phenology module is part of a DGVM that is applied for climate impact studies. In order to assess how well the model performs under different climate conditions, we additionally tested how the model performance changes in grid cells that were not used during optimization (Figure 7.26). We found no general decrease in model performance with distance to the nearest grid cell used for optimization, or under different temperature conditions. Especially, no significantly lower correlations ( $p \leq 0.05$ , Wilcoxon rank-sum test, Figure 7.26) were found between simulated and observed FAPAR time series in grid cells that were 3 to 5°C warmer than the closest optimization grid cell. From a typical perspective of space for time substitution, this could indicate that the confidence in the simulation of FAPAR dynamics should not detract under climate warming scenarios of 0.3 to 4.8°C (IPCC, 2014).

Nevertheless, model optimization experiments and model evaluation indicated further needs for improvement in future studies – for example, simulations of surface albedo could improve through the implementation of time-varying effects of snow conditions and surface moisture on albedo. Also, an enhanced representation of canopy architecture and canopy radiative transfer could reduce the large spatial variability and parameter uncertainty found for the light extinction coefficient and hence improve the simulation of tree coverage and peak FAPAR. In addition to temperature, light and water availability, phenology also depends on other factors that are not considered in LPJmL-GSI. Phenology is also driven by leaf age (Caldararu et al., 2012, 2014) and nutrient availability (Wright, 1996). These effects are neither considered in the original GSI phenology model (Jolly et al., 2005; Stöckli et al., 2011) nor in the LPJmL-GSI or other traditional formulations. Here, the lower posterior values found for the parameter  $\alpha_a$  may be compensating for missing nitrogen limitation effects on productivity in boreal forests (Vitousek and Howarth, 1991). Thus a future implementation of nitrogen limitation processes in LPJmL requires a re-optimization of the  $\alpha_a$  parameter. Additionally, the current implementation of phenology in LPJmL affects photosynthesis only through changes in APAR. In future model developments a

stronger coupling between phenology and ecosystem carbon cycle dynamics could be explored. For example, the LPJmL-GSI phenology module could demand carbon for leaf development from photosynthesis or additional storage pools on the one hand and could trigger carbon turnover through litterfall on the other hand. In this case a phenology module could partly regulate an optimal carbon gain for a canopy similar to the approach of Caldararu et al. (2014). Nevertheless, such an analysis needs to go beyond the approach of Caldararu et al. (2014) and demands for additional observational constraints on ecosystem carbon fluxes, leaf area, biomass and litterfall. In order to better understand couplings between leaf phenology, changes in carbon allocation and photosynthesis it will be of benefit to use site level eddy covariance measurements from the FLUXNET network (Baldocchi et al., 2001) together with ancillary data in ecosystem-scale model optimization experiments (Carvalhais et al., 2010; Kuppel et al., 2012; Williams et al., 2009). Thus the LPJmL-GSI phenology module and the LPJmL model–data integration approach can serve as a framework to further explore hypotheses of ecosystem processes and vegetation dynamics.

We demonstrated the improved performance of LPJmL-GSI over LPJmL-OP in representing observed carbon fluxes and stocks, forest cover and seasonal to decadal dynamics of vegetation greenness. Thus, similar approaches to the LPJmL-GSI phenology module can be applied in other DGVMs to improve model simulations in comparison with observations. However, the adaptation of current results to other models should be cautionary because the phenology scheme of LPJmL-GSI is an empirical approach with PFT-dependent parameters that need to be estimated. This estimation is model-specific because (1) different models do not necessarily use the same definition and set of PFTs; (2) our parameterizations depend on model structure, e.g. different models often use different hydrology routines; and (3) our posterior parameters for phenology were also constrained by using albedo and GPP data. Thus LPJmL-GSI model parameters cannot be easily transferred to other models. It might be possible to use the parameters of the temperature and light limiting functions in other models because these functions depend uniquely on the forcing data. On the other hand, the parameters for the water availability limiting function might need to be re-optimized because of differences in soil moisture computations. However, depending on the co-variability between forcing variables and simulated water availability by

other models, the best parameterizations may differ from the ones presented here. Consequently, we acknowledge the potential need to optimize parameters of the LPJmL-GSI phenology model in order to obtain plausible results in other modelling structures. However, it is likely that the LPJmL-GSI phenology model can be easily applied to other models of the LPJ group of models (Prentice et al., 2011; Smith et al., 2001) that are using the hydrology routines of Gerten et al. (2004) while probably additional parameter optimization exercise are required to adapt the model to other types of DGVMs or ecosystem models.

#### **3.3.4 Environmental controls on vegetation greenness phenology**

As the newly developed GSI-based phenology model of LPJmL can reproduce the seasonality and monthly dynamics of observed FAPAR in most biomes, it can be used to identify environmental controls on vegetation greenness phenology. The importance of phenological controls differed by climate regions, ecosystems and season (Figure 3.9). We identified environmental controls on seasonal FAPAR dynamics by analysing the mean seasonal cycles of FAPAR, of the cold temperature, light, water availability and heat stress limiting functions for phenology from the LPJmL-GSI model run. This analysis is comparable to previous investigations of limiting factors for vegetation phenology (Caldararu et al., 2014; Jolly et al., 2005). FAPAR seasonality in high-latitude regions (tundra, boreal forests) was mainly controlled by cold temperature (entire year) and light (October to February). We also found an important control by water availability in February to April in the tundra and in boreal forests of North America and eastern Siberia. This water limitation in early spring was due to the seasonal freezing of the upper permafrost layer in LPJmL. FAPAR seasonality in temperate grasslands in western North America and central Asia was controlled from a mixture of cold temperature (January to April), of water availability (May to November) and light (November to January). FAPAR seasonality in temperate forests in Europe was mainly limited by cold temperature in spring and by a combination of cold temperature and light in autumn. Additionally, heat stress and water availability contributed to a small reduction in summer FAPAR in temperate and boreal forests. The FAPAR seasonality in savannas (Sahel) was limited by water availability in the entire year and additionally by heat stress before the beginning of the rain season. The

FAPAR seasonality of temperate regions in South America was limited by water availability in the entire year. Cold temperature was additionally limiting between May and September. Thus, water availability was the only environmental factor in LPJmL-GSI that controlled phenology globally from tropical to Arctic biomes.

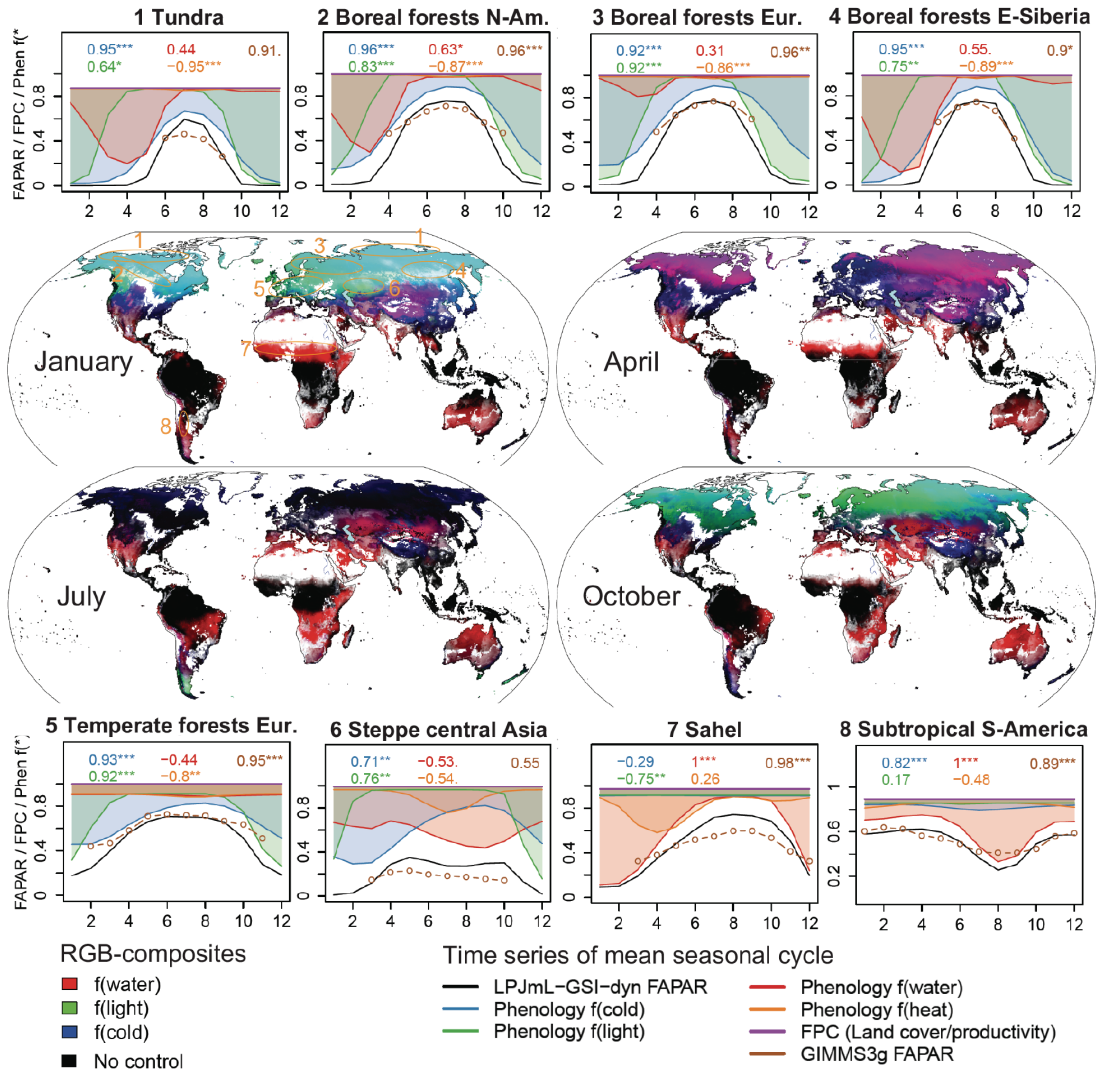


Figure 3.9: Phenological controls on seasonal FAPAR dynamics.

The maps are red–green–blue composites of the mean monthly values for the water (red), light (green) and cold temperature (blue) phenology limiting function values from the LPJmL-GSI model run. White regions in the maps are without vegetation or dominated by croplands for which the LPJmL-GSI phenology model was not applied. Time series represent the mean seasonal cycles (January to December) (averaged over 1982–2011) of simulated and observed FAPAR and phenology limiting function values averaged for different regions as indicated in the first map. Phenology limiting function values close to 0 indicate a strong control by phenology limiting functions whereas values close to 1 indicate no phenological control. The correlation coefficients of each time series with the simulated FAPAR time series are shown in each time series plot. The significance of the correlation is indicated as point symbol (see Figure 3.7 for an explanation of significance symbols).

The implementation of the water limiting function on phenology in LPJmL-GSI resulted in unique patterns of phenological controls that were different from results reported in similar analyses (Caldararu et al., 2014; Jolly et al., 2005). LPJmL-GSI showed water limitation on phenology in many subtropical and dry temperate regions (especially Mediterranean, Pampas and Patagonia in South America, Mongolia, and northern Great Plains). The original GSI model showed mainly temperature and light limitation in these regions. In contrast to the original GSI, our implementation considers water limitations on phenology based on plant available water and not on VPD (Jolly et al., 2005). As considered by Caldararu et al. (2014), soil water availability exerts a more direct control on phenology development, which has been demonstrated for Mediterranean ecosystems (Kramer et al., 2000; Richardson et al., 2013) and in dry temperate grasslands (Liu et al., 2013; Yuan et al., 2007).

Additionally, we identify water availability as an important limiting function for spring phenology in boreal and Arctic regions in LPJmL-GSI because of the seasonal freezing of the upper active layer in permafrost soils. Although no relationships between active layer depth and vegetation greenness were found so far (Mcmichael et al., 1997), frozen grounds limit the seasonal tree growth in boreal forests because of limited water supply and nutrient uptake (Benninghoff, 1952; Jarvis and Linder, 2000). As the seasonal freezing and thawing of permafrost soils is to a large extent driven by changes in air temperature, one might argue that air temperature is enough to explain phenology dynamics in boreal and Arctic regions. Nevertheless, we found weak correlations between posterior model parameters for the cold temperature and water limiting function for phenology in PFTs that experience strong permafrost dynamics (BoNS  $r = 0.2$ , PoH  $r = -0.28$ ) (Figure 7.16). This indicates that the water and cold temperature limiting functions in boreal and Arctic regions are only weakly correlated. Indeed, we did not find a completely synchronized temporal dynamic of the cold temperature and water limiting functions for phenology (Figure 3.9). These results emphasize the ability to disentangle effects of seasonal air temperature and soil moisture on phenology in boreal and Arctic regions. Air temperature and soil thawing are not completely synchronized because soil temperature depends also on topography, substrate and the insulating effects of the snow, litter and vegetation cover (Jorgenson et al., 2010; Shur and Jorgenson, 2007; Zhang, 2005). Soils might be still

frozen if air temperature is already positive or vice versa. Also experimental studies highlighted the role of permafrost-regulated soil moisture on phenology and productivity in boreal and Arctic ecosystems (Natali et al., 2012; Schuur et al., 2007). It also has been observed that the seasonal freezing and thawing in permafrost regions regulates ecosystem evapotranspiration (Ohta et al., 2008) and fire activity (Forkel et al., 2012) especially during extreme dry years. Thus, although temperature might be enough to explain average spatial patterns of phenology in boreal and Arctic regions we acknowledge that variations in snow or vegetation cover that affects soil temperature and thus moisture might be important factors in explaining inter-annual variations of land surface phenology.

The heat stress limiting function was additionally introduced in LPJmL-GSI. Heat stress had no importance for seasonal FAPAR dynamics in most regions except in temperate and tropical grasslands. The heat stress function was highly correlated with the water availability function in temperate grasslands. This suggests that summer FAPAR is both regulated by water-induced and temperature-induced drought conditions in temperate grasslands. In tropical grasslands, heat stress and water availability were driving the temporal dynamics of seasonal FAPAR but asynchronously (in the Sahel). These results suggest that soil moisture needs to be considered in observational data analyses and in other ecosystem models as a controlling factor for vegetation phenology in all biomes.

Interestingly, Caldararu et al. (2014) identify leaf age as the dominant factor for phenology development in many permanent moist subtropical and tropical forests, but also in several water-limited regions which were here identified as seasonally controlled by water availability. We cannot identify a dominant control on seasonal FAPAR dynamics in evergreen forests, as leaf age is not explicitly simulated in LPJmL-GSI. We acknowledge that the consideration of leaf age effects on phenology could further enhance the representation of ecosystem processes. However, the seasonal co-variation between LAI or FAPAR and environmental controls on phenology complicates the ability to disentangle the leaf aging signal from a temperature, light or water availability-driven signal, especially in seasonally deciduous vegetation types, where climate-driven models explain a significant fraction of seasonal variability and the realized age of leaves is shorter than a year. In addition, cloud cover contamination

over moist tropical or subtropical forests pertain usually a weak seasonal signal and a high short-term variability, hinging on the reliability of the seasonal signal. In particular, Morton et al. (2014) show that seasonal changes in MODIS LAI in the Amazon forests are linked to insufficient corrections of the sun–sensor geometry, which challenge the representation of vegetation phenology. However, in these tropical moist regions, where we find no environmental seasonal controls, and the realized age of oldest leafs are higher than a year, leaf age may be an important contributor for further consideration regarding the above-seasonal frequency of phenology. Hence, grasping the relevance of leaf longevity, especially in tropical perennial systems, would necessarily require ground observations of leaf development and litter fall to constrain leaf age parameters, as well as measurements of soil water content to address the appropriateness of soil moisture effects.

### **3.4 Conclusions**

We have demonstrated a major improvement of the LPJmL dynamic global vegetation model by implementing a new set of phenological controls on vegetation greenness and by integrating multiple decadal satellite observations. We have proven that the original phenology model in LPJmL is unable to explain temporal dynamics of FAPAR. As an alternative we implemented a new phenology model (LPJmL-GSI) which considers effects of cold temperature, heat stress, light, and water availability on vegetation phenology. We developed a model–data integration approach for LPJmL (LPJmL-MDI) to (1) constrain model parameters against observations, (2) to directly integrate observed land cover fractions and burnt area time series and (3) to evaluate LPJmL against independent data streams. Specifically, phenology, productivity, and albedo-related model parameters of LPJmL-GSI were optimized jointly against 30-year time series of satellite observations of FAPAR, against 10-year time series of vegetation albedo and against mean annual patterns of gross primary production using a genetic optimization algorithm.

The new phenology model and the parameter optimization clearly improved LPJmL model simulations. LPJmL-GSI better reproduces observed spatial patterns of gross primary production, tree cover, biomass and FAPAR than the original model. LPJmL-



GSI simulates global total carbon stocks and fluxes that are closer to independent estimates than from the original model. LPJmL-GSI better represents observed seasonal, monthly, inter-annual and decadal FAPAR dynamics than the original model. The improvements of LPJmL in representing observed patterns and temporal dynamics of vegetation greenness allows assessing environmental controls on vegetation phenology and greenness. Contrasting to previous studies (Jolly et al., 2005; Stöckli et al., 2011), our results indicate that soil water availability is a major control of seasonal FAPAR dynamics not only in water-limited biomes but also in boreal forests and the Arctic tundra where water availability is regulated through seasonal thawing and freezing of the active permafrost layer. Until now the phenology of these ecosystems was mostly considered as temperature-limited. The consideration of the effect of soil water availability on phenology in LPJmL improved model simulations of greening trends in the Sahel and of browning trends in parts of the boreal forests of North America. Our results demonstrate that improved phenology models that consider seasonal effects of water availability on a continuous canopy development are needed in order to correctly explain seasonal to long-term dynamics in vegetation greenness.

## **Acknowledgements**

We thank Maarten Braakhekke and Enrico Tomelleri for testing model optimization algorithms. We thank Ulrich Weber for his assistance in data processing. We thank Ranga Myneni and Frédéric Baret for their comments on FAPAR data sets. We thank the IT departments at MPI-BGC Jena and PIK Potsdam for providing and maintaining the respective high performance computing infrastructures.

We gratefully thank the following researchers, groups and institutes for producing, providing and hosting their data sets:

- Compton Tucker, Jorge Pinzon, Ranga Myneni and the GIMMS group for the GIMMS FPAR3g data set
- INRA, CNES and VITO for the Geoland2 FAPAR data set
- Martin Jung for the MTE GPP and ET data, and the SYNMAP land cover map

- Sassan Saatchi for the tropical forest biomass map – Markus Kottek, Franz Rubel and the University of Veterinary Medicine Vienna for providing the Koeppen-Geiger climate map
- Louis Giglio and Guido Van der Werf for providing the GFED database
- Jennifer L. Northway and Gary Schmunk for compiling the Alaska Large Fire Database
- Natural Resources Canada for providing the Canada National Fire Database
- Phil Jones, Ian Harris and the CRU research unit for providing the CRU climate data set
- ECMWF for ERA-Interim climate reanalysis data.

M. Forkel received funding from the Max Planck Institute for Biogeochemistry and from the European Commission's 7th Framework Programme project CARBONES (grant agreement 242316). M. Forkel conducted this work under the International Max Planck Research School for Global Biogeochemical Cycles.

The service charges for this open access publication have been covered by the Max Planck Society.



## **4 Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness**

---

This chapter has been submitted as:

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U. and Carvalhais, N.: Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness, submitted to *Global Change Biology*, 2014-12-16.

Supporting Information of this chapter can be found in chapter 8.

## 4 Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness

### Abstract

Identifying the relative importance of climatic and other environmental controls on the inter-annual variability and trends in global land surface phenology and greenness is challenging. Firstly, quantifications of land surface phenology and greenness dynamics are impaired by differences between satellite datasets and phenology detection methods. Secondly, dynamic global vegetation models (DGVM) that can be used to diagnose controls still reveal structural limitations and contrasting sensitivities to environmental drivers. Thus we assessed the performance of a new developed phenology module within the LPJmL (Lund Potsdam Jena managed Lands) DGVM with a comprehensive ensemble of three satellite datasets of vegetation greenness and ten phenology detection methods, thereby thoroughly accounting for observational uncertainties. The improved and tested model allows us quantifying the relative importance of environmental controls on inter-annual variability and trends of land surface phenology and greenness at regional and global scales. We found that start of growing season inter-annual variability and trends are in addition to cold temperature mainly controlled by incoming radiation and water availability in temperate and boreal forests. Warming-induced prolongations of the growing season in high latitudes are dampened by a limited availability of light. For peak greenness, inter-annual variability and trends are dominantly controlled by water availability and land use and land cover change (LULCC) in all regions. Stronger greening trends in boreal forests of Siberia than in North America are associated to a stronger increase in water availability from melting permafrost soils. Our findings emphasize that in addition to cold temperatures, water availability is a co-dominant control for start of growing season and peak greenness trends at the global scale.

## 4.1 Introduction

Satellite observations demonstrated globally significant inter-annual variability and trends of phenology and greenness in the last three decades (Myneni et al., 1997a; Tucker et al., 2001; Xu et al., 2013; Zeng et al., 2013). The use of satellite-derived time series of vegetation indices, such as the normalized difference vegetation index (NDVI) (Tucker, 1979) to study the timing of changes in vegetation greenness is usually referred as land surface phenology (de Beurs and Henebry, 2004a). Variability and trends in land surface phenology and greenness have been associated with regionally different climatic and environmental controlling factors: Positive trends in land surface greenness (“greening”) (Goetz et al., 2005; Myneni et al., 1997a) and phenological changes like an earlier start and a lengthening of the growing season (Julien and Sobrino, 2009; Tateishi and Ebata, 2004; Tucker et al., 2001) in high-latitude regions have been concordantly associated to warming climate (Keenan et al., 2014; Lucht et al., 2002; Menzel et al., 2006; Xu et al., 2013). Nevertheless, also increasing atmospheric CO<sub>2</sub> and nitrogen deposition can potentially fertilize vegetation and thus contribute to global greening trends (Mao et al., 2012; Piao et al., 2006). The CO<sub>2</sub> fertilization effect is supposed to be mainly important in drylands (Donohue et al., 2013). On the other hand, multiple controls have been identified for negative trends in vegetation greenness (“browning”) in boreal forests of North America, and in some temperate and subtropical grasslands (Bi et al., 2013; Goetz et al., 2005; de Jong et al., 2013b) like fire regimes (Goetz et al., 2005), heat stress (Bunn et al., 2007), forest type (Beck and Goetz, 2011), cooling spring temperatures (Wang et al., 2011), reduced soil moisture, and possibly permafrost (Barichivich et al., 2014). In the Sahel, greening trends are discussed in face of opposing effects of increasing precipitation and increasing land degradation (Dardel et al., 2014a; Fensholt et al., 2013). Land management might be also important for end of growing season in temperate regions (Garonna et al., 2014). For tropical forests, it has been intensively discussed whether vegetation index time series have seasonal dynamics and if these are driven by variations in light or water availability (Huete et al., 2006; Morton et al., 2014; Samanta et al., 2010). Seasonal to decadal dynamics of land surface greenness affect ecosystem structure (Fridley, 2012; Wolkovich and Cleland, 2010) and the climate system through changes in albedo, surface roughness, and through exchange of energy, water and carbon (Bonan, 2008b;

Richardson et al., 2013). Consequently, it is important to understand the relative importance of regional competitive explanations and controlling factors on average spatial patterns, inter-annual variability and trends in land surface phenology and greenness. An observation-based identification of this relative importance is difficult because of several challenges that limit the use of explanatory data analysis approaches because controlling factors: 1) act on different spatial scales (e.g. uniform atmospheric CO<sub>2</sub> increase vs. regional fire events), 2) are temporally correlated (e.g. CO<sub>2</sub> and air temperature increase), 3) exhibit non-linear dynamic interactions (e.g. drought, fire, land cover change and succession), or 4) are not readily available from observations (e.g. spatial distributed observation of permafrost dynamics). On the other hand, a consistent framework like a dynamic global vegetation model (DGVM) is not limited by these challenges if such processes are accurately represented in the model and can be applied to identify the relative importance of controlling factors like temperature, prescription, fire disturbance, CO<sub>2</sub> fertilization, permafrost dynamics and soil moisture, and land use and land cover change (LULCC) (Piao et al., 2011).

DGVMs were previously applied to identify controlling factors for land surface phenology and greenness (Lucht et al., 2002; Mao et al., 2012, 2013; Piao et al., 2006) but need to be critically evaluated with respect to the model performance in reproducing observations. Especially, DGVMs cannot well reproduce observed phenology (Richardson et al., 2012) and seasonal to decadal dynamics of land surface greenness (Anav et al., 2013; Murray-Tortarolo et al., 2013). On the other hand, new phenology models have been recently developed and parameterized using satellite-derived greenness observations (Caldararu et al., 2014; Forkel et al., 2014; Knorr et al., 2010; Stöckli et al., 2011). This new generation of phenology models describes the temporal development of canopy greenness and thus follows a different paradigm than traditional phenology models that usually simulate specific events of leaf development like budburst or leaf senescence (Richardson et al., 2012). Forkel et al. (2014) developed and parameterized a new phenology model within the LPJmL (Lund Potsdam Jena managed Lands) DGVM that better reproduces satellite observations of seasonal to decadal dynamics of the fraction of absorbed photosynthetic active radiation (FAPAR) by considering effects of temperature, light and water availability. Using LPJmL with this improved phenology scheme and additional model developments (Bondeau et al.,

2007; Schaphoff et al., 2013; Thonicke et al., 2010) allows going beyond an earlier LPJ-based analysis of climate effects on peak greenness (Lucht et al., 2002) to additionally quantify effects of water availability, light, fire, permafrost, and land cover dynamics on land surface phenology and greenness.

The estimation of land surface phenology and greenness metrics (PGMs) and their inter-annual and decadal dynamics from satellite-derived time series is challenging for several reasons: 1) PGMs like the start and end of the growing season (SOS, EOS) are known to control the annual CO<sub>2</sub> uptake but they are limited descriptors of site-specific seasonal changes in canopy structure and plant physiology (Fu et al., 2014a; Liang et al., 2011; Schwartz and Reed, 1999; Studer et al., 2007; Xiao et al., 2009). 2) Satellite datasets differ remarkably in spatial patterns and temporal dynamics due to different sensor properties (sun-sensor geometry, spectral and spatial resolution), observational distortions (cloud and snow cover, aerosols) and processing algorithms (D'Odorico et al., 2014; Fensholt and Proud, 2012; Guay et al., 2014; Jiang et al., 2013; McCallum et al., 2010; Scheffic et al., 2014; Wang et al., 2012). Additionally, the temporal resolution of satellite datasets and thus of greenness time series affects the timing of phenological or trend changes but it has been shown that temporal resolution is of minor importance for the variability in timing in comparison to the statistical time series analysis method (Forkel et al., 2013; White et al., 2014; Zhang et al., 2009). Thus, 3) the estimation of PGMs from vegetation index time series is highly sensitive to the chosen analysis methods. Satellite-derived greenness time series require usually smoothing and interpolation to exclude short-term variability and to estimate daily phenology events from less frequent observations. Therefore different smoothing, interpolation, curve fitting and detection methods were developed (hereinafter simplifying called "phenology methods") (de Beurs and Henebry, 2010b). These methods can result in remarkable differences in estimated patterns and dynamics of land surface phenology and greenness (Forkel et al., 2013; Geng et al., 2014; Hird and McDermid, 2009; Kandasamy et al., 2013; Musial et al., 2011; White et al., 2014). Thus, it is necessary to consider the uncertainty from different datasets and phenology methods for a robust quantification and for model evaluation of recent patterns and temporal dynamics of land surface phenology and greenness.



Here we aim to 1) to quantify the relative importance of climatic and ecosystem controlling factors on average patterns, inter-annual variability and trends in land surface phenology and greenness at the regional and global scale; for which we need to 2) quantify global patterns and temporal dynamics of land surface phenology and greenness by considering uncertainties from different satellite datasets and phenology methods; and to 3) evaluate the performance of the LPJmL DGVM with an improved phenology scheme (Forkel et al., 2014) in reproducing the observed patterns and dynamics when considering these uncertainties.

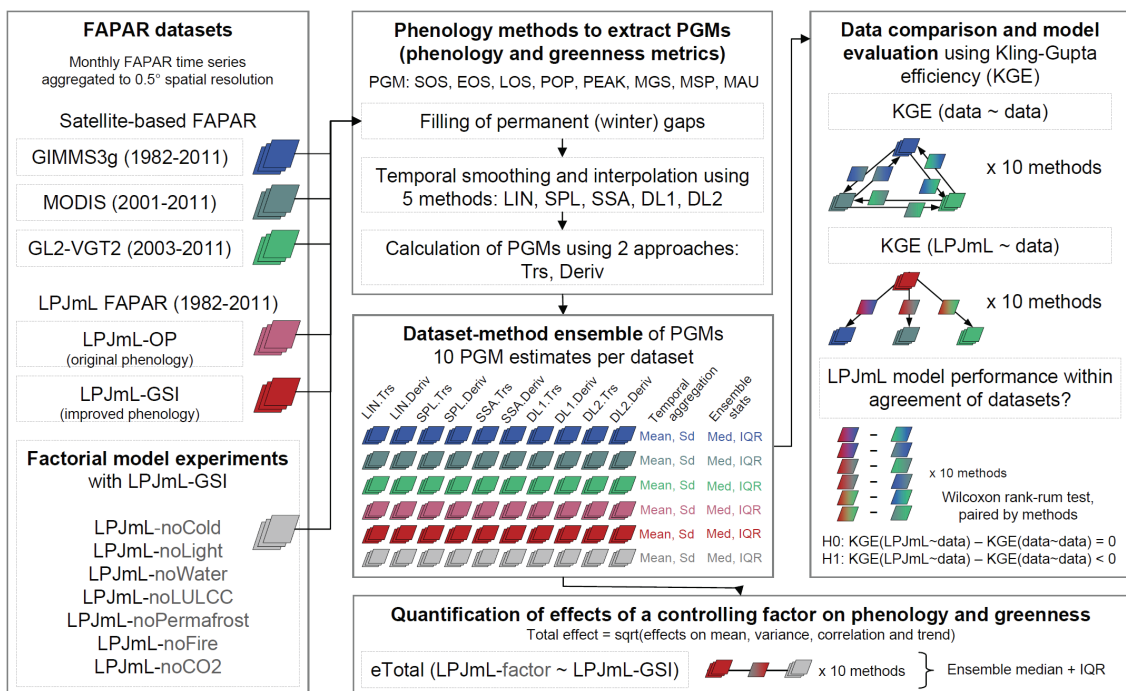


Figure 4.1: Flow chart of the described datasets and methods.

Abbreviations of methods are explained in the Data and methods section and in Table 8.1, respectively.

## 4.2 Material and methods

### 4.2.1 FAPAR datasets

FAPAR time series used here originate from three satellite-based datasets and from LPJmL with two different phenology schemes (Figure 4.1). The GIMMS3g (Global Inventory Modeling and Mapping Studies, 3rd generation) FAPAR dataset was derived from a harmonized NDVI dataset (Pinzon and Tucker, 2014; Zhu et al., 2013) and covers fully the years 1982 to 2011. MODIS FAPAR (Moderate-Resolution Imaging

Spectroradiometer) was taken from the MOD15A2 product (Knyazikhin et al., 1999; USGS, 2001) and covers the period since February 2000. GL2-VGT2 FAPAR (Geoland2 BioPar GEOV1, Vegetation 2) was derived from SPOT (Satellite Pour l'Observation de la Terre) observations (Baret et al., 2013). The dataset covers originally the period 1999 to 2012 based on VGT1 and VGT2 observations. However, we only use the dataset from 2003 onwards (only VGT2 observations) because the combined dataset has discontinuities at the sensor shift between VGT1 and VGT2 (Forkel et al., 2014; Horion et al., 2014). FAPAR from GIMMS3g, MODIS and GL2-VGT2 are defined as black-sky green canopy instantaneous FAPAR at 10:35, 10:35 and 10:15 solar time, respectively. Instantaneous FAPAR observations during this time are close approximations of the daily integrated FAPAR (Baret et al., 2007) and thus comparable with LPJmL model simulations. All satellite-based datasets were aggregated from their original resolution to 0.5° spatial resolution and to monthly time steps to be comparable with LPJmL model simulations. Sub-monthly FAPAR values were aggregated to monthly values using the maximum value composite approach (Holben, 1986) to exclude potential remaining low-biased FAPAR values.

#### **4.2.2 FAPAR and phenology in LPJmL**

The LPJmL DGVM simulates carbon and water fluxes and stocks as a result of natural vegetation dynamics (Sitch et al., 2003) and has extensions for human land use and agriculture (Bondeau et al., 2007), fire (Thonicke et al., 2010), surface albedo and snow cover (Strengers et al., 2010), and soil hydrology and permafrost dynamics (Schaphoff et al., 2013). Additionally, a new phenology scheme for natural vegetation plant functional types (PFT) has been implemented in LPJmL and FAPAR-related model parameters were optimized against GIMMS3g FAPAR, albedo and gross primary production data (Forkel et al., 2014). FAPAR of a PFT in LPJmL depends mainly on annual changing foliar projective cover (FPC) and the daily phenology status (PHEN) (Forkel et al., 2014).

We applied two phenology schemes in LPJmL to simulate PHEN. LPJmL-OP is the original LPJ phenology scheme (Sitch et al., 2003) and is based on accumulated temperature conditions (i.e. growing-degree days) for summergreen PFTs. LPJmL-GSI (growing season index) is an alternative and improved phenology scheme and

simulates seasonal leaf development in response to cold temperature, light, water availability and heat stress controlling functions:

$$PHEN_{PFT} = f_{cold,PFT}(T) \times f_{light,PFT}(SW_{down}) \times f_{water,PFT}(W) \times f_{heat,PFT}(T) \quad (4.1)$$

where T is the average daily air temperature, SW<sub>down</sub> the daily short-wave radiation and W is the percentage water availability (Forkel et al., 2014).

#### 4.2.3 LPJmL model setup and factorial model experiments

LPJmL was driven by monthly time series of air temperature and precipitation from the CRU TS3.1 dataset (Harris et al., 2014), and by monthly short-wave downward and long-wave net radiation time series from the ERA-Interim reanalysis dataset (Dee et al., 2011). Monthly observed burnt area time series were prescribed to the fire module as described in Forkel et al. (2014) to constrain fire simulation by observations. For this, we used burnt area estimates from the Global Fire Emissions Database (GFED4) (Giglio et al., 2010) for the period 1996-2011, from the Alaskan Large Fire Database (Frames, 2012; Kasischke et al., 2002), and from the Canadian National Fire Database (CFS, 2010; Stocks et al., 2002) for North America for the period 1979-1996. To assess LULCC effects, we prescribed to LPJmL a dataset of recent and historic (1700-2005) cropland distributions (Fader et al., 2010). All model simulations and the required model spinup were performed according to the standard LPJmL modelling protocol (Schaphoff et al., 2013; Thonicke et al., 2010). Transient model runs were analyzed for the period 1982-2011 for which FAPAR satellite observations are available.

We performed several experiments with LPJmL-GSI to quantify the effects of different controlling factors on inter-annual variability and trends in land surface phenology and greenness (Table 4.1). Specifically, we investigated the effects of seasonal climatic controls (cold temperature, light and water availability), fire, land use and land cover change (LULCC), permafrost dynamics and CO<sub>2</sub> fertilization by running the LPJmL-GSI phenology model. In a second step, we run a series of model experiments using the same set-up and drivers as the standard model run but with one factor fixed at a time (Table 4.1). We fixed the cold temperature, light, and water availability controlling functions of the LPJmL-GSI phenology scheme to unity to assess the effects of temperature, light and water availability, respectively. This implies that temperature, light or water availability do not affect FAPAR phenology in LPJmL-GSI but can still

affect productivity and thus FAPAR through annual changes in FPC. All other model experiments directly affected FAPAR dynamics through annual changes in FPC. We disabled the simulation of soil thermal dynamics to assess a possible effect of permafrost. This experiment implies that seasonal soil freezing and thawing does not affect soil moisture and rooting depth and thus generally increases productivity and thus FPC. We made another model run without simulating any fire activity to assess the possible effect of fire disturbance. To assess the importance of LULCC, we fixed land use and land cover fractions. For this, cropland fractions and the maximum FPC of natural vegetation in a grid cell were kept constant after 1982. This approach implies that the FPC of PFTs can still change because of competition or mortality but the area extent cannot exceed the coverage conditions of 1982. To assess the effect of CO<sub>2</sub> fertilization, atmospheric CO<sub>2</sub> fractions were kept constant at 341.22 ppm after 1982, and thus did not further fertilize photosynthesis and increase FPC. Finally, we performed one model experiment with the original LPJmL phenology model (LPJmL-OP) (Sitch et al., 2003) to use a classical growing degree day-based phenology model as an alternative and benchmark in model evaluation.

Table 4.1: Overview of factorial model experiments and corresponding effects on FAPAR in LPJmL.

LPJmL model experiment	Factor	Effect on FAPAR in LPJmL	Phenology	Factorial changes to LPJmL model components				
				Permafrost	Agriculture (land use)	Natural vegetation	Burnt area	CO <sub>2</sub>
LPJmL-GSI	Standard	--	GSI	yes	yes	dynamic	observed	growing
LPJmL-OP	--	--	OP	yes	yes	dynamic	observed	growing
LPJmL-noCold	Cold	Direct effects on phenology status	GSI, but $f_{cold} = 1$	yes	yes	dynamic	observed	growing
LPJmL-noLight	Light		GSI, but $f_{light} = 1$	yes	yes	dynamic	observed	growing
LPJmL-noWater	Water	PHEN (daily)	GSI, but $f_{water} = 1$	yes	yes	dynamic	observed	growing
LPJmL-noLULCC	LULCC	Direct effect on FPC (annual)	GSI	yes	land use fractions fixed to 1982	maximum FPC fixed to 1982	observed	growing
LPJmL-noPf	Permafrost (Pf)	Indirect effects on FPC (annual)	GSI	no	yes	dynamic	observed	growing
LPJmL-noFire	Fire		GSI	yes	yes	dynamic	no fire	growing
LPJmL-noCO2	CO <sub>2</sub>		GSI	yes	yes	dynamic	observed	constant after 1982

#### 4.2.4 Phenology methods and trend analysis

Different phenology methods were used to estimate the uncertainty in PGMs that is caused by different smoothing, interpolation and analysis methods (Figure 4.1).

Although FAPAR time series from LPJmL model simulations are gap-free and without observational distortions like in satellite datasets, we applied all phenology methods also to modelled FAPAR time series to ensure comparability with estimated PGMs from satellite datasets. All applied methods are freely available in the R software package “greenbrown” (<http://greenbrown.r-forge.r-project.org/>).

In our approach, all phenology methods consist of three steps (Figure 4.1): 1) permanent gaps (i.e. usually winter months in northern regions) were filled in each time series; 2) the time series were smoothed and interpolated to daily time steps using five different methods; 3) PGMs were calculated from smoothed and daily interpolated time series using two different approaches. All methods are described with more details within the Supporting Information 1.

In the first step, we filled months with permanent gaps series (i.e. gaps that occur in at least 20% of all years during the same season) with the minimum FAPAR value. This approach was already used by Beck et al. (2006) to fill missing winter observation in NDVI time series.

In the second step, we used five different methods for temporal smoothing and for interpolation to daily values (Supporting Information 1). These methods use linear interpolation, spline smoothing and interpolation (Migliavacca et al., 2011; Musial et al., 2011), singular spectrum analysis (Golyandina et al., 2001; Mahecha et al., 2010a), or two curve-fitting approaches with double-logistic functions (Beck et al., 2006; Elmore et al., 2012) to derive daily interpolated and smoothed FAPAR time series.

In the third step, we used the smoothed and daily interpolated time series to estimate start of growing season (SOS) and end of growing season (EOS) by either using 50% thresholds on the seasonal greenness curve (approach Trs) (White et al., 1997) or the derivative of the seasonal curve (approach Deriv) (Tateishi and Ebata, 2004) (Supporting Information 1). Both approaches are based on the definition of SOS and EOS as the midpoints of spring green-up and autumn senescence, respectively. We followed this definition of SOS and EOS (Tateishi and Ebata, 2004; White et al., 1997), although lower thresholds or extreme values of the second derivative of the seasonal greenness curve better agree with phenology transitions observed at the surface (White et al., 2014). Nevertheless, SOS and EOS definitions that are based on lower values are more strongly affected from non-vegetation changes as snow cover or cloud

contaminations and thus less reliable. All other PGMs were derived afterwards: The length of the growing season (LOS) is the difference between EOS and SOS. Mean growing season FAPAR (MGS) is the average FAPAR value from all days between SOS and EOS. Mean spring (MSP) and mean autumn (MAU) FAPAR are the average FAPAR values from a period of 20 days around SOS and EOS, respectively. Peak FAPAR (PEAK) is the maximum FAPAR value of the year from the smoothed and interpolated curve. The position of the peak (POP) is the day of the year when PEAK is reached. In summary we used ten phenology methods (five smoothing and interpolation methods in step #2 times two detection approaches in step #3). Thus we derive for each dataset and each PGM an ensemble of ten annual time series of land surface phenology and greenness metrics.

Trends in all annual PGM time series were computed based on linear least-square regression with breakpoint detection (Bai and Perron, 2003; Forkel et al., 2013; Zeileis et al., 2003). The significance of the trend was estimated by using the Mann-Kendall trend test (Kendall, 1975; Mann, 1945).

#### 4.2.5 Data comparison, model evaluation and quantification of factorial effects

We compared the time series of PGM derived from all satellite datasets to assess their agreement and their uncertainty with the aim of evaluating the performance and usability of LPJmL. Specifically we assessed the agreement of PGM time series regarding the mean, two measures of inter-annual variability (standard deviation and correlation) and overall agreement. For this we computed the Kling-Gupta efficiency (KGE) with its components that account for bias, difference in standard deviation and correlation (Gupta et al., 2009). KGE ranges between negative infinity (worst agreement) and 1 (perfect agreement) and is defined based on the Euclidean distance in a 3-dimensional coordinate system of agreement measures:

$$KGE = 1 - \sqrt{(\alpha - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \quad (4.2)$$

where  $\gamma$  is the Pearson correlation coefficient between two time series.  $\alpha$  and  $\beta$  account for the difference in standard deviation  $\sigma$  and in the mean values  $\mu$  between a times series  $x$  and a reference time series  $r$ , respectively:

$$\alpha = \frac{\sigma_x}{\sigma_r} \quad (4.3)$$

$$\beta = \frac{\mu_x}{\mu_r} \quad (4.4)$$

We computed KGE between two datasets for the same phenology method. This resulted for each satellite dataset in an ensemble of 20 KGEs (e.g. GIMMS3g compared against MODIS and GL2-VGT2 using 10 methods each) and for each LPJmL model experiment in an ensemble of 30 KGEs (i.e. LPJmL compared against 3 satellite datasets using 10 methods each). We tested if the performance of LPJmL is within the agreement of the datasets by testing if the differences between KGE from LPJmL against satellite datasets and KGE from the cross-comparison of satellite datasets are significant less than 0 by using the Wilcoxon rank-sum test (paired along phenology methods, Figure 4.1).

We also applied the KGE metric to quantify the effect of a factor on PGM time series in the factorial model experiment. For this we computed KGE between a PGM time series from a factorial model run and the reference LPJmL-GSI model run for each phenology method. In order to additionally quantify the effect on the trend, we extended the KGE metric by a fourth metric  $\delta$  which accounts for the differences in linear trend slopes  $\tau$  between a times series from a factorial model experiment  $x$  and a time series  $r$  from the reference model run:

$$\delta = \frac{\tau_x}{\tau_r} \quad (4.5)$$

Thus, we are defining the total effect ( $eTotal$ ), the effect on the mean ( $eMean$ ), the effect on the variance of annual values ( $eVar$ ), the effect on correlation or inter-annual dynamic ( $eCor$ ), and the effect on the trend ( $eTrend$ ) of a factor accordingly to the Kling-Gupta efficiency as the Euclidean distance in a 4-dimensional space:

$$eTotal = \sqrt{eVar + eMean + eCor + eTrend} \quad (4.6)$$

where  $eVar$ ,  $eMean$ ,  $eCor$  and  $eTrend$  are defined as the squared differences between 1 and  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$ , respectively. To consider the uncertainty of phenology methods, we are reporting results as the median and inter-quartile range of an effect over all phenology methods.

## 4.3 Results

### 4.3.1 Comparison of land surface phenology and greenness metrics from datasets and models

Spatial patterns of phenology and greenness metrics differed remarkably between datasets (Figure 4.2, Table 4.2). In most biomes, the mean annual SOS and EOS dates were detected earlier in the GIMMS3g dataset and later in the GL2-VGT2 dataset (Figure 4.2, Table 4.2). Mean annual SOS differed between satellite datasets by up to 50 days in Savannas and in boreal needle-leaved summergreen forests. Mean annual SOS and EOS dates simulated by LPJmL-OP were out of phase in temperate grasslands and in Savannas. All satellite datasets and LPJmL-GSI agreed regarding the global patterns of mean annual peak FAPAR but LPJmL-OP overestimated peak FAPAR in temperate to arctic regions (Figure 4.2 c). These differences between LPJmL-OP and LPJmL-GSI are related to structural limitations hampering the optimization of LPJmL-OP (Forkel et al., 2014).

Estimated PGMs had large differences between phenology methods (Figure 4.2, Table 4.2). For example, mean annual SOS ranged over almost 80 days for the GL2-VGT2 dataset in boreal needle-leaved summergreen forests. These differences cannot be solely associated to distortions of optical remote sensing observations but also to weaknesses on the phenology methods as well, since similar differences were found for the LPJmL model. These wide ranges cannot be related to single methods because the estimated PGMs of a method depend also on biome and dataset, i.e. a method might result in an extreme SOS in one biome or dataset whereas it might result in an average SOS in another biome or dataset (Figure 8.5). Thus the use of a single phenology method can result in wrong conclusions about land surface phenology and greenness dynamics in some biomes.



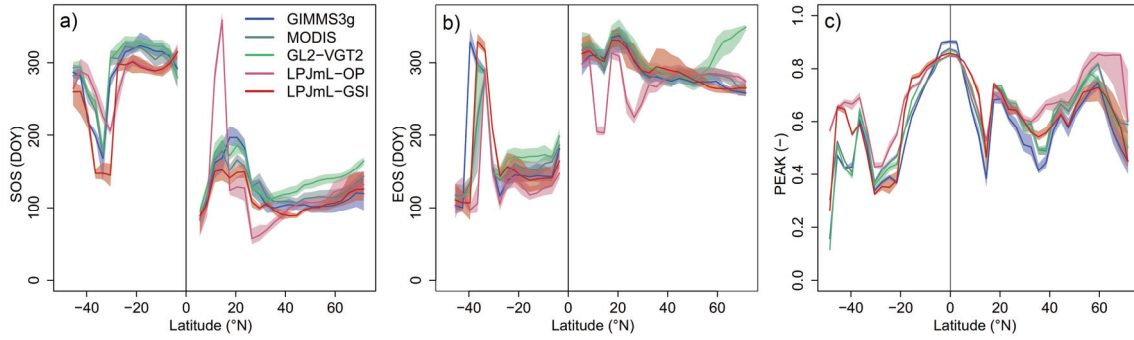


Figure 4.2: Latitudinal gradients of mean annual (a) SOS, (b) EOS and (c) peak FAPAR from datasets and LPJmL.

For each dataset or model the median and the inter-quartile range of the method ensemble are shown.

Table 4.2: Biome-averaged mean annual start of season (SOS) and end of season (EOS) (in days of year, DOY) from satellite-based datasets (GIMMS3g, MODIS, GL2-VGT2) and LPJmL phenology models.

Numbers are the mean, standard deviation and the minimum and maximum values (in square brackets) of SOS (or EOS) from 10 phenology methods. Star symbols indicate the p-value of a two-sided Wilcoxon rank-sum test (paired by phenology method) if the multi-method ensemble of SOS (or EOS) estimates of a dataset or model equals the ensemble of all other satellite-based datasets (null hypothesis) or if it is outside the dataset ensemble (alternative hypothesis). P-values are: \*\*\*  $p \leq 0.001$ , \*\*  $0.001 < p \leq 0.01$ , \*  $0.01 < p \leq 0.05$ , no symbol for  $p > 0.05$ . All biome-averaged values were derived for the northern hemisphere (except Savannas). Results in **bold font** highlight datasets or models without a significant difference to the other datasets. See Figure 8.1a for definitions of biomes.

Biome	PGM	GIMMS3g	MODIS	GL2-VGT2	LPJmL-OP	LPJmL-GSI
Savannas (N-hemisphere)	SOS	137±8 [129, 153] ***	146±10 [137, 165] **	147±8 [137, 161] **	<b>133±20 [105, 157]</b>	<b>138±5 [130, 146]</b>
	EOS	<b>287±9 [270, 301]</b>	<b>286±12 [270, 306]</b>	291±14 [269, 318] *	272±20 [253, 298] ***	280±7 [267, 290] ***
Savannas (S-hemisphere)	SOS	261±13 [233, 275] ***	277±9 [261, 292] **	<b>275±12 [253, 293]</b>	291±8 [277, 305] ***	265±10 [246, 273] *
	EOS	152±6 [143, 161] ***	<b>161±8 [151, 176]</b>	170±9 [158, 185] ***	139±18 [119, 167] ***	<b>162±14 [146, 183]</b>
Temperate grasslands	SOS	113±6 [106, 123] ***	<b>120±6 [112, 132]</b>	139±9 [124, 150] ***	102±5 [98, 112] ***	<b>129±7 [119, 138]</b>
	EOS	266±12 [247, 290] **	<b>272±15 [253, 302]</b>	<b>271±12 [252, 290]</b>	289±9 [280, 306] ***	258±14 [244, 281] ***
Temperate broad-leaved summergreen forests	SOS	98±4 [92, 104] ***	<b>104±5 [98, 111]</b>	118±6 [111, 127] ***	88±4 [84, 97] ***	87±5 [77, 95] ***
	EOS	288±13 [274, 305] ***	<b>294±13 [281, 312]</b>	300±12 [284, 320] ***	280±14 [264, 301] ***	<b>292±17 [275, 316]</b>
Boreal needle-leaved evergreen forests	SOS	100±5 [94, 106] ***	<b>113±9 [102, 131]</b>	140±6 [129, 146] ***	<b>118±8 [108, 132]</b>	105±9 [92, 120] **
	EOS	286±14 [274, 313] ***	<b>295±15 [280, 325]</b>	306±16 [262, 320] *	<b>292±11 [277, 306]</b>	268±10 [251, 279] ***
Boreal needle-leaved summergreen forests	SOS	107±16 [85, 129] ***	<b>121±11 [101, 132]</b>	149±22 [120, 199] ***	136±24 [102, 180] *	<b>116±16 [94, 137]</b>
	EOS	257±10 [237, 268] **	<b>272±11 [255, 284]</b>	<b>270±27 [212, 301]</b>	275±9 [258, 287] **	<b>260±12 [239, 276]</b>
Tundra	SOS	120±21 [90, 147] ***	<b>138±14 [112, 160]</b>	160±8 [149, 178] ***	<b>142±21 [113, 175]</b>	123±16 [99, 146] ***
	EOS	263±10 [244, 272] ***	<b>276±12 [259, 287]</b>	291±24 [228, 317] **	<b>282±10 [263, 296]</b>	<b>269±12 [247, 282]</b>

Table 4.3: Biome-averaged trends in SOS (days/year).

Numbers are the mean, standard deviation and the minimum and maximum values (in square brackets) of SOS trend slopes from 10 phenology detection methods. Numbers in round brackets indicate the number of methods that resulted in significant SOS trends ( $p \leq 0.05$ , Mann-Kendall trend test). Results in **bold font** highlight trends for which most methods agree in trend direction and for which at least one method indicated a significant trend. See Figure 8.1a for definitions of biomes.

Biome	GIMMS3g	MODIS	GL2-VGT2	LPJmL-OP	LPJmL-GSI
Time period	1982-2011	2001-2011	2003-2011	1982-2011	1982-2011
Savannas: N-hemisphere	0.04±0.16 [-0.24, 0.26] (3*)	-0.35±0.33 [-0.96, 0.15] (1*)	0.36±0.4 [0.03, 1.07] (0*)	0.13±0.14 [-0.09, 0.41] (5*)	<b>0.09±0.04</b> <b>[0.01, 0.14] (1*)</b>
Savannas: S-hemisphere	-0.03±0.1 [-0.18, 0.12] (2*)	0.19±0.54 [-0.67, 0.99] (1*)	<b>0.64±0.6</b> <b>[-0.33, 1.56]</b> <b>(1*)</b>	<b>0.21±0.09</b> <b>[0.08, 0.39] (7*)</b>	0.08±0.19 [-0.21, 0.39] (4*)
Temperate grasslands	<b>-0.07±0.04</b> <b>[-0.13, 0] (2*)</b>	0.39±0.22 [0.1, 0.72] (0*)	0.5±0.16 [0.19, 0.76] (0*)	<b>-0.1±0.07</b> <b>[-0.19, 0.01]</b> <b>(6*)</b>	<b>-0.1±0.07</b> <b>[-0.22, -0.02]</b> <b>(2*)</b>
Temperate broad-leaved summer- green forests	-0.12±0.49 [-1.22, 1.31] (7*)	0±0.13 [-0.22, 0.18] (0*)	<b>-0.32±0.27</b> <b>[-0.94, 0.16]</b> <b>(1*)</b>	-0.09±0.04 [-0.12, -0.01] (0*)	<b>-0.09±0.05</b> <b>[-0.16, -0.01]</b> <b>(1*)</b>
Boreal needle- leaved evergreen forests	<b>-0.11±0.05</b> <b>[-0.23, -0.06]</b> <b>(4*)</b>	-0.29±0.26 [-0.88, 0.14] (0*)	-0.2±0.29 [-0.48, 0.31] (0*)	<b>-0.12±0.05</b> <b>[-0.2, -0.01]</b> <b>(5*)</b>	<b>-0.11±0.04</b> <b>[-0.18, -0.05]</b> <b>(5*)</b>
Boreal needle- leaved summergreen forests	-0.03±0.05 [-0.11, 0.04] (1*)	<b>-0.94±0.25</b> <b>[-1.33, -0.49]</b> <b>(7*)</b>	0.09±1.63 [-1.31, 3.74] (1*)	<b>-0.14±0.11</b> <b>[-0.29, 0.08]</b> <b>(4*)</b>	<b>-0.23±0.13</b> <b>[-0.53, -0.09]</b> <b>(5*)</b>
Tundra	-0.01±0.05 [-0.09, 0.07] (0*)	<b>-0.27±0.17</b> <b>[-0.66, 0.02]</b> <b>(3*)</b>	-0.1±0.18 [-0.33, 0.19] (0*)	<b>-0.17±0.08</b> <b>[-0.26, -0.03]</b> <b>(7*)</b>	<b>-0.18±0.08</b> <b>[-0.3, -0.07]</b> <b>(8*)</b>

Temporal dynamics and trends in PGMs agreed only partly between datasets (Figure 4.3, Table 4.3). Peak FAPAR had significant greening in tundra, boreal forests of Siberia, and temperate forests of Europe in GIMMS3g and LPJmL-GSI in 1982-2011. In boreal forests of North America, peak FAPAR had significant greening in LPJmL-GSI and negative but non-significant trends in the satellite-based datasets. Nevertheless, in the overlapping period of all datasets (2003-2011), GIMMS3g had positive peak FAPAR trends whereas MODIS and GL2-VGT2 had negative trends, reflecting also a disparity between datasets for this region. In the Sahel, most datasets had greening trends (but only for GIMMS3g significant), which was reproduced by LPJmL-GSI. In the Amazon, only the GIMMS3g dataset had significant greening whereas MODIS, GL2-VGT2 and LPJmL-GSI had no trends. In the Congo basin, only MODIS had significant browning whereas the other datasets had positive trends (GL2-VGT2) or no trends (GIMMS3g, LPJmL-GSI). All datasets had (partly significant) trends towards earlier SOS in boreal

and temperate forests and in the tundra (Figure 4.3 h-i). Negative SOS trends in temperate forests of Europe were in agreement with positive trends in LOS (Figure 8.3). We found no trends in LOS in boreal forests of Siberia because trends towards earlier SOS were compensated by trends towards an earlier EOS (Figure 8.3). These parallel changes in spring and autumn phenology have also been observed in some temperate tree species (Fu et al., 2014c). LPJmL-GSI reproduced the observed SOS, EOS and LOS trends in these regions. We found no significant trends in SOS, EOS or LOS in other regions (Figure 8.3). Nevertheless, SOS trends for one dataset had partly large uncertainties and even opposite trend directions depending on the phenology method (Table 4.3). These results demonstrate the need to use multiple phenology methods and satellite datasets in order to robustly provide confidence for the application of LPJmL to identify controlling factors for land surface phenology and greenness.

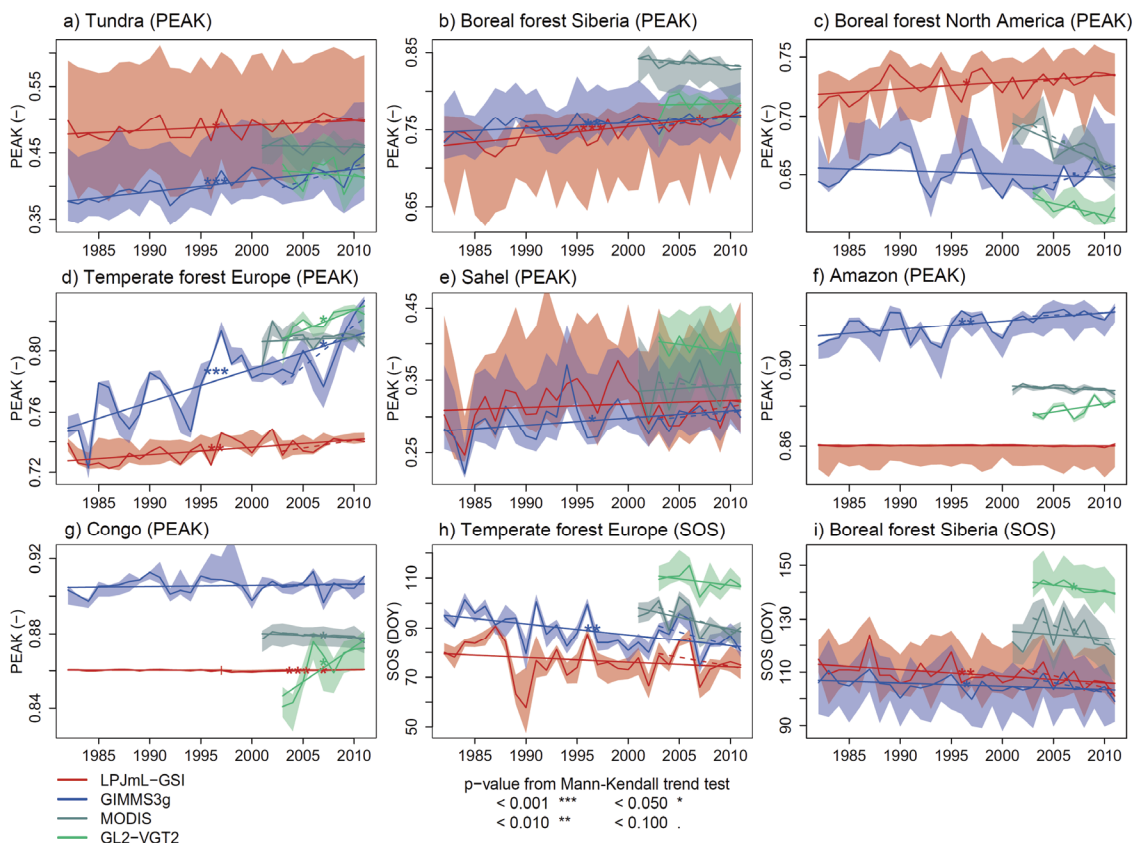


Figure 4.3: Regional averaged time series and linear trends of (a-g) peak FAPAR and (h-i) SOS. Time series are the multi-method median and inter-quartile range. Dashed lines indicate the linear trend for the overlapping period of all datasets (2003-2011). The significance of the trend in each segment is indicated by star symbols. See Figure 8.1b for definitions of regions.

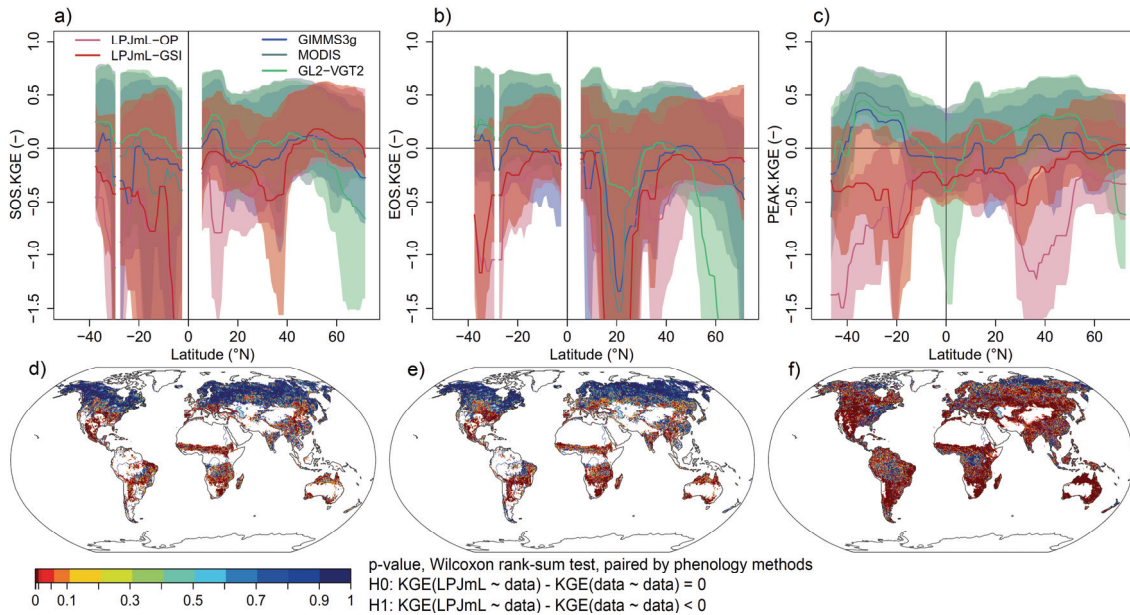


Figure 4.4: Kling-Gupta efficiency for SOS, EOS and peak FAPAR for datasets and LPJmL phenology models.

(a-c) Latitudinal gradients of KGE. For each dataset the median KGE and the range between the 1st quartile and the maximum of the method ensemble are shown. As KGE can potentially reach until negative infinity, the figures are cutted at -1.5. (d-e) p-value maps for the test if the performance of LPJmL-GSI is within the agreement of datasets.

The evaluation of the performance of LPJmL-OP and LPJmL-GSI with respect to the uncertainty of satellite datasets and phenology methods showed that LPJmL-GSI agreed for SOS and EOS in temperate, boreal and arctic regions better with the satellite-based datasets than the datasets with each other (Figure 4.4, Figure 4.5 i-k). Especially in boreal forests and tundra, the KGE of LPJmL-GSI for SOS, EOS and LOS was in more than 96% of all grid cells within the KGE ensemble of the satellite datasets. In regions south of 40°N, the KGE ensemble of LPJmL was usually significantly lower than the KGE ensemble of each dataset. Nevertheless, it is possible to find always a combination of a satellite dataset and phenology method that attests the model either a modest or almost perfect performance for SOS (Figure 8.7). For peak FAPAR, LPJmL-GSI had lower KGEs than the satellite data ensemble in most regions except the arctic tundra and tropical evergreen forests (Figure 4.4) where satellite datasets had generally the lowest agreements (Figure 4.5 a-h). LPJmL-GSI had the weakest performance for peak FAPAR in comparison to the satellite datasets in the Sahel where the KGE was only in 6% of all grid cells within the KGE ensemble of the datasets. In all other regions, the performance of LPJmL-GSI in reproducing peak FAPAR was in at least 29% of all grid cells within the agreement of the datasets. Globally, the null hypothesis

that the KGE ensemble for peak FAPAR and SOS from LPJmL-GSI equals the KGE ensemble of dataset cross-comparisons was accepted in 42% and 74% of all land grid cells, respectively. The KGE of LPJmL-GSI was for most PGMs and grid cells higher than the KGE of LPJmL-OP. Low KGEs between satellite datasets and between LPJmL-GSI and satellite datasets were usually related to weak correlations and secondly to differences in standard deviation (Figure 4.5). In summary, we demonstrate large uncertainties in the timing of land surface phenology from different methods and satellite-based datasets but also highlight the improved usability of LPJmL-GSI over LPJmL-OP to explain phenology and greenness dynamics.

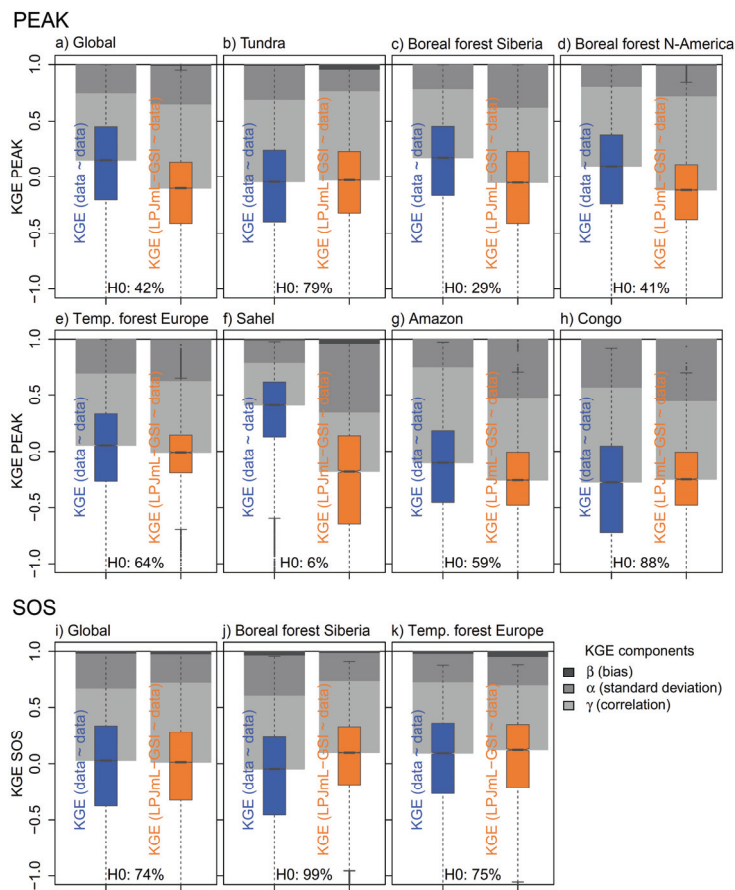


Figure 4.5: Kling-Gupta efficiency (KGE) for satellite data comparison and LPJmL-GSI model evaluation for peak FAPAR (a-h) and SOS (i-k) globally and for different regions. Boxplots are regional distributions of KGE from all grid cells of a region and from the full data-method ensemble. Plots are limited to -1. Barplots show the regionally averaged (median) components of KGE and are plotted as the difference between the regional median KGE and a perfect KGE at 1. H0 indicates the percentage of grid cells in a region for which the null hypothesis,  $H_0: KGE(LPJmL-GSI \sim data) - KGE(data \sim data) = 0$ , was accepted (one-sided Wilcoxon rank-sum test, paired by phenology methods,  $p \leq 0.05$ ).

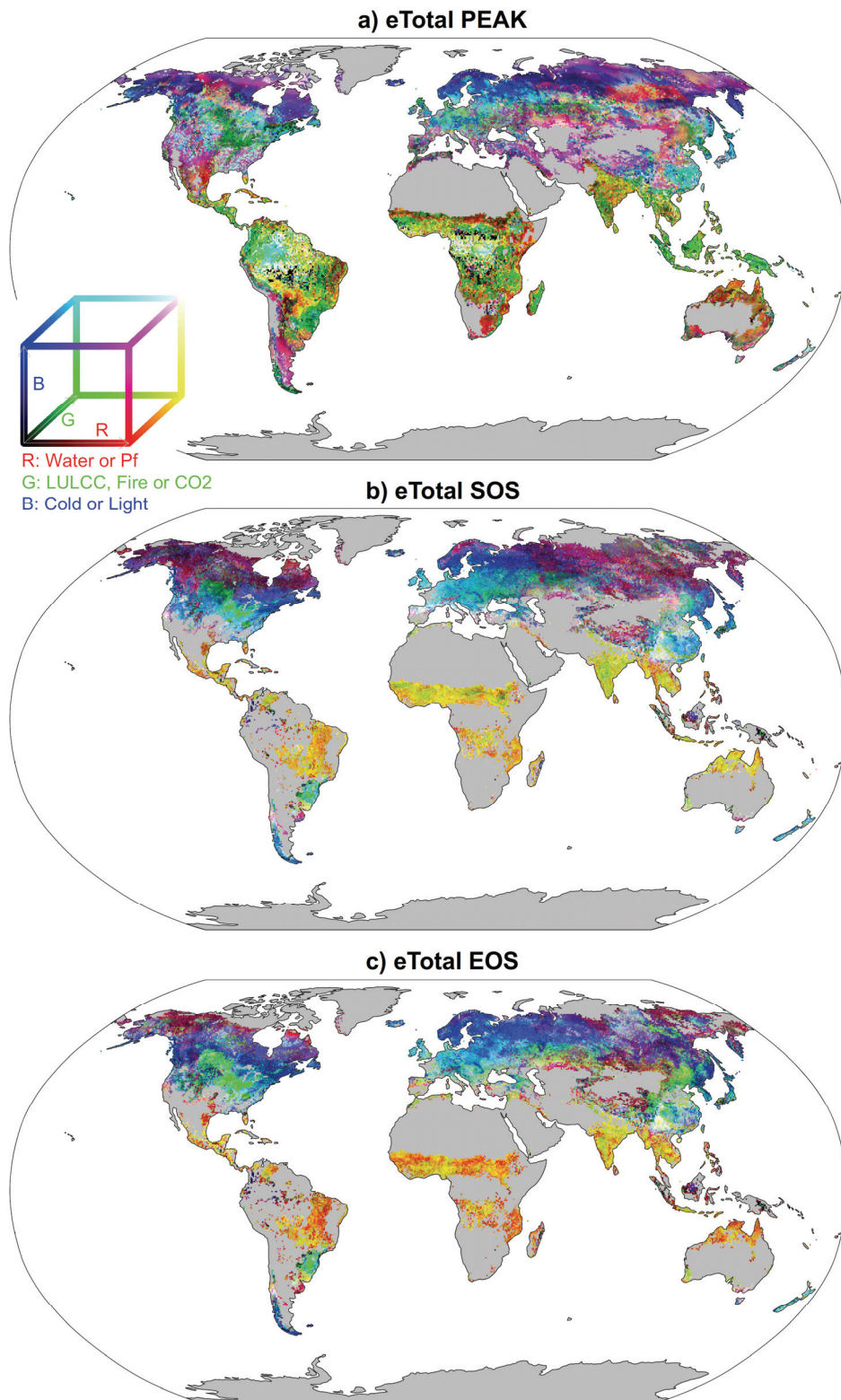


Figure 4.6: Total effect of different factors on (a) peak FAPAR, (b) SOS and (c) EOS. The maps are red-green-blue composites of the maximum eTotal of grouped factors. Red colours indicate a high total effect of seasonal water limitation or permafrost, green colours of LULCC, fire or CO<sub>2</sub>, and blue colours of seasonal cold or light limitation. In case of black or white colours all factors have a low or high total effect, respectively. Shown are the eTotal results from the median of the phenology method ensemble. All eTotal values were mean-centered, scaled and histogram stretching was applied to enhance plotting. Grey areas are without vegetation or seasonality was not detected.

### 4.3.2 Controls on phenology and greenness dynamics

We identified seasonal water limitation and land-use driven changes in vegetation composition as the globally dominant controls on annual changes in vegetation phenology and greenness (Figure 4.6, Figure 8.6). Water availability had the highest total effects on peak FAPAR in latitudes south of 40°N (Figure 4.7). Seasonal cold temperature and light limitation determined total effects on peak FAPAR in regions pole-wards of 30°N and 30°S, respectively. A multitude of factors had high total effects on peak FAPAR in temperate, boreal and arctic regions north of 40°N, especially through seasonal cold temperature, seasonal light limitation, LULCC, and permafrost. CO<sub>2</sub> fertilization had an effect on peak FAPAR in all regions but was of minor importance in comparison to all other factors. Fire had small total effects on peak FAPAR at the global scale but dominated regionally in fire-prone ecosystems such as Savannas and parts of boreal forests.

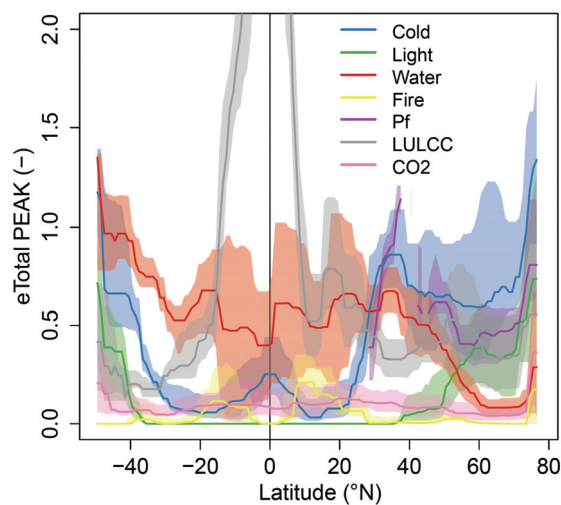


Figure 4.7: Latitudinal gradients of the total effect of different factors on peak FAPAR. Shown is the median and inter-quartile range of the phenology method ensemble. eTotal values were averaged (median value) over 1° bands and a 7° running median filter was applied.

The inter-annual variability of land surface phenology was not only affected by seasonal effects of cold temperature, light and water availability, but also by annual effects like LULCC and fire (Figure 4.6 b-c). In subtropical regions (South America, Sahel, India, northern Australia), water availability together with LULCC and fire had high total effects on SOS, whereas EOS was more influenced by water availability. In

temperate regions (eastern US, Europe), SOS and EOS were mostly affected by seasonal cold temperature, seasonal light limitation and by LULCC. In agricultural regions in the central US, Ukraine and southern Russia, LULCC was the dominant controlling factor for SOS and EOS. In boreal and arctic regions, a mixture of seasonal cold temperature, light, water availability and permafrost dynamics had high total effects on SOS and EOS. These results demonstrate the importance of water availability, LULCC and permafrost dynamics for annual dynamics in vegetation phenology.

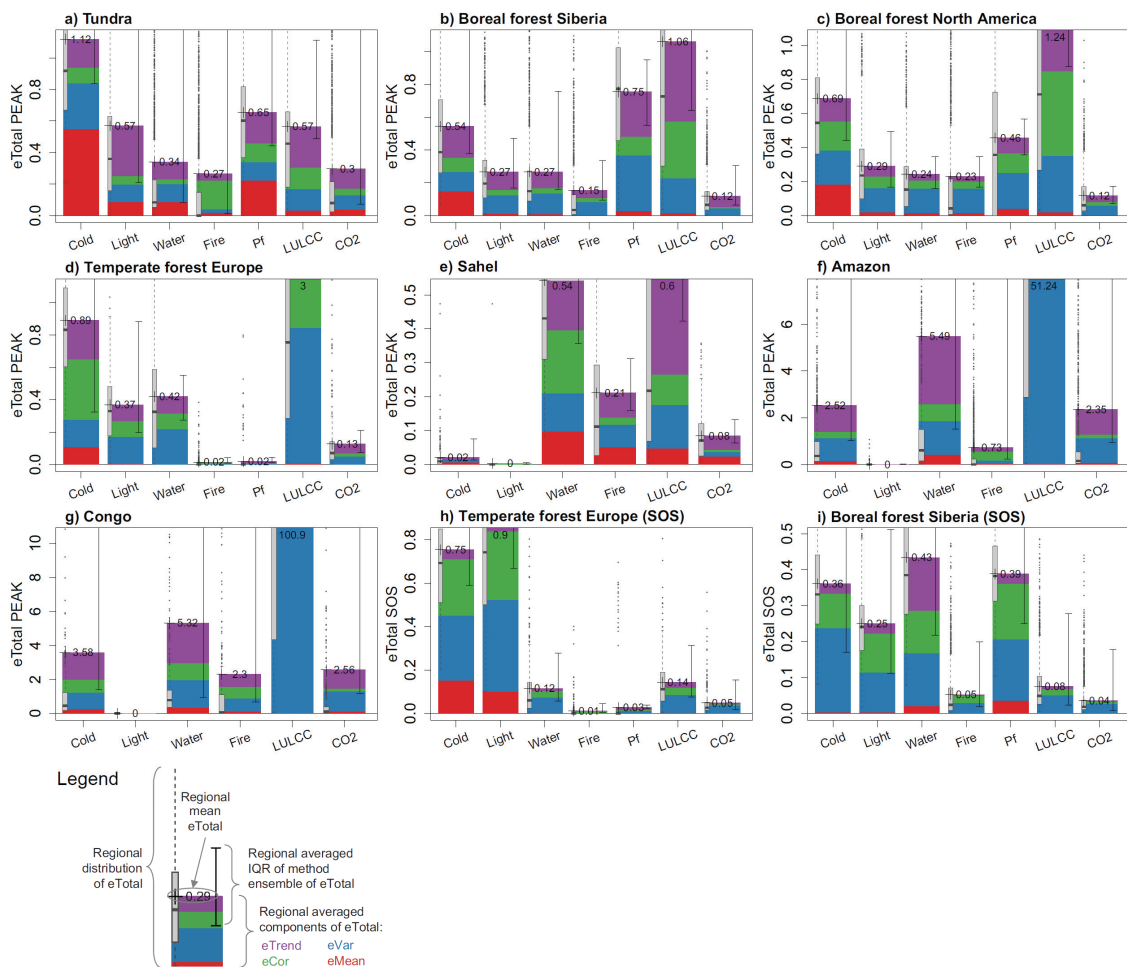


Figure 4.8: Regional distributions and averaged total effects with components on (a-g) peak FAPAR and (h-i) SOS.

Environmental factors had regionally different effects on mean, variance, correlation and trends of vegetation phenology and greenness (Figure 4.8). In the tundra, the variance and correlation of annual peak FAPAR was mainly affected by seasonal cold temperature, permafrost and LULCC (Figure 4.8 a). Light limitation strongly influenced the trend in peak FAPAR because LPjM-L-GSI resulted in even stronger



peak FAPAR greening trends if the length of the growing season in tundra regions would not be limited by radiation. In boreal forests of Siberia, permafrost determined the variance in annual peak FAPAR whereas both LULCC and permafrost had high effects on trends in peak FAPAR (Figure 4.8 b). On the other hand, permafrost dynamics were less important in boreal forests of North America than in Siberia: The variance of annual peak FAPAR was equally affected by a mixture of seasonal cold, light and water limitation whereas LULCC had the highest effect on correlation and trends (Figure 4.8 c). In temperate forests of Europe, peak FAPAR was mostly affected by LULCC, which mainly affected the variance, correlation and trend of annual dynamics (Figure 4.8 d). In the Sahel, seasonal water availability and LULCC had both high effects on the mean, variance, correlation and trends in peak FAPAR (Figure 4.8 e). In tropical forest regions, the trends in peak FAPAR were dominantly affected by water availability (Figure 4.8 f, g). Although the total effect of CO<sub>2</sub> fertilization was small in all regions, CO<sub>2</sub> affected mainly trends in peak FAPAR.

In temperate forests of Europe the SOS was mainly affected by seasonal cold temperature and light limitation (Figure 4.8 h). In boreal forests of Siberia SOS trends were additionally affected by permafrost (Figure 4.8 i) pointing to a combined effect of increasing temperatures, melting permafrost and thus higher water availability. Light had an additional strong effect on SOS trends because these trends would be even stronger if light is not ultimately limiting the growing season in these high-latitude regions.

On the global scale, two satellite datasets and the LPJmL-GSI standard simulation had greening trends and trends towards earlier start of season (but not significant) (Figure 4.9 a-b). These positive trends in peak FAPAR and negative trends in SOS are well correlated with global trends in air temperature and water availability (Figure 4.9 c-d). Positive trends in water availability from LPJmL-GSI are related to positive trends in global precipitation (Figure 8.8). Interestingly, peak FAPAR from LPJmL-GSI and GL2-VGT2, and SOS from LPJmL-GSI, GIMMS3g and MODIS had globally stronger correlations with water availability than with temperature (Figure 4.9 b). Specifically, the trends and variance of global peak FAPAR was mostly affected by seasonal cold temperature and water limitation and LULCC in the factorial model experiment (Figure 4.9 e). Correlation and trends in SOS were in the global average dominated by

seasonal cold temperature and water limitation (Figure 4.9 f). These results suggest that increasing water availability is a co-dominating control for the global greening and start of season trends, along with air temperature.

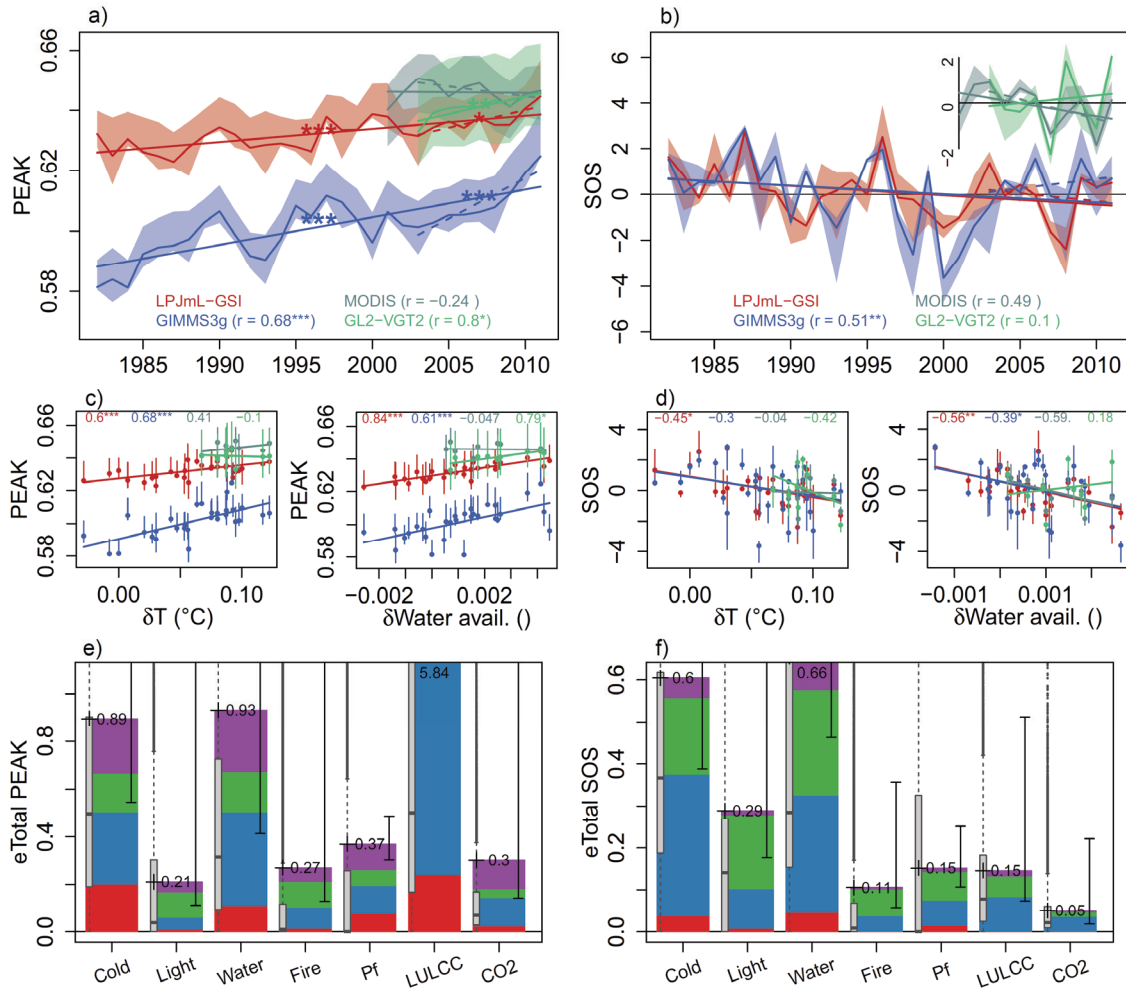


Figure 4.9: Global temporal dynamic and controlling factors for land surface phenology and greenness.

Globally averaged time series with trends in (a) peak FAPAR and (b) SOS anomalies, respectively. SOS anomalies are differences to the mean SOS of each grid cell. SOS time series from MODIS and GL2-VGT2 were shifted by +4 days to improve the readability. Scatterplots with linear regression and correlation coefficients between (c) peak FAPAR and (d) SOS, and global averaged anomalies (relative to 1982) of air temperature (CRU TS3.2) and water availability (LPJmL) for satellite datasets and LPJmL-GSI. Global distribution and global averaged total effects with components on (e) peak FAPAR and (f) SOS. See Figure 8 for an explanation of (e) and (f).

## 4.4 Discussion

### 4.4.1 Uncertainties in detection and explanation of land surface phenology and greenness

Our results demonstrated the need to use multiple datasets and phenology methods for detection and model evaluation of land surface phenology and greenness metrics. Moreover, the detection and explanation of land surface phenology and greenness dynamics is affected from several sources of uncertainties like: 1) the used observational response variable (here FAPAR) (Walker et al., 2014; White et al., 2014; Wu et al., 2014), 2) the definition of phenology events, 3) the phenology method applied, 4) the used observational datasets with associated errors and differences in processing algorithms, and 5) the explanatory model.

We did not account for the uncertainty sources #1 and #2 in this study to ensure the comparability of LPJmL model results with satellite-based datasets. A major source of uncertainty or mismatches between different PGM estimates is the definition of phenology events (especially SOS and EOS) (de Beurs and Henebry, 2010b). Commonly, SOS and EOS from vegetation index time series are either defined as (A) the start of spring greenup and the end of autumn senescence (Jonsson and Eklundh, 2002; Zhang et al., 2004) or (B) as the mid-days of spring greenup and autumn senescence (Karlsen et al., 2006; Tateishi and Ebata, 2004; White et al., 1997). The latter definition was used in this study. Definition (A) has the disadvantage that such observations are often affected by soil reflectance, snow cover or are prove to noise (Beck et al., 2006; Delbart et al., 2006; Huete et al., 1992) and thus might result in a wrong timing. These are less important issues for definition (B). Although our results by using definition (B) are to a limited extent comparable with ecological changes at the surface (White et al., 2014), we avoid drawing conclusions about dynamics or model performance that were likely affected by such observational limitations.

The use of different phenology methods (uncertainty #3) is one of the major source of uncertainty (de Beurs and Henebry, 2010b). Smoothing and interpolation methods have been developed to exclude short-term variability and non-vegetation changes from phenology detection (Beck et al., 2006; Delbart et al., 2006; Jonsson and Eklundh, 2002). The accuracy of these methods often depends on the number of missing observations, land cover type, dataset properties and the magnitude of short-term and

inter-annual variability (Forkel et al., 2013; Geng et al., 2014; Kandasamy et al., 2013; Musial et al., 2011; Verbesselt et al., 2010a). Additionally, the approach to detect phenology dates (i.e. thresholds, extreme values of the derivative of the smoothed time series, or parameters of fitting functions) introduces further differences (de Beurs and Henebry, 2010b; White et al., 2014). It is possible to identify an optimal phenology method for an ecosystem if ground observations are available (White et al., 2014). Nevertheless, for continental to global applications an identification of an optimal phenology method is not straightforward. Phenology methods tuned in a particular area with ground observation can be too specific and not easily generalizable and usable in another area. This implies that a phenology method might result in a good agreement between datasets whereas another phenology method might result in a worse agreement (Fig. S4e). Thus, the use of a single method may result in erroneous diagnostics on phenology dynamics, and may undermine the robustness in analysing dataset agreement or model performance. Unless a single robust method can be identified from ground observations, the use of ensembles of state of the art methods for detection of land surface phenology and greenness provide an empirical probabilistic approach to detect PGMs.

The fourth source of uncertainty arises from the variety of datasets with differences in sensor properties, observational errors and processing algorithms. Our results confirm previous findings of negative SOS trends in Europe, North America and Asia (Hamunyela et al., 2013; Julien and Sobrino, 2009; Zhang et al., 2014) and of greening trends from the GIMMS3g dataset (Bi et al., 2013; Fensholt and Proud, 2012; de Jong et al., 2013b; Xu et al., 2013). The weak agreement between datasets was more related to low correlations and variance than to biases (Figure 4.5). Low correlations between PGMs were found previously in Europe (Atzberger et al., 2013), Mongolia (Miao et al., 2013), tundra, temperate arid regions and tropical forests (Fensholt and Proud, 2012), and in parts of the USA and Mexico (Scheftic et al., 2014). Given these differences between datasets, the use of a single dataset for model evaluation might result in biased conclusions about model performance (Anav et al., 2013; Murray-Tortarolo et al., 2013). Thus, the use of multiple datasets (Guay et al., 2014) and their uncertainty is required for a robust analysis and of land surface phenology and greenness dynamics.

Finally, the structure, parameters and forcing data of the used model to explain land surface phenology and greenness is another uncertainty source (Migliavacca et al., 2012; Raupach et al., 2005). Differences between forcing datasets like precipitation introduce a high model uncertainty in the simulation of FAPAR inter-annual variability especially in regions with sparse coverage of weather stations (Traore et al., 2014). Thus the low performance of LPJmL in the Sahel is probably to a large extent caused by the uncertainty in the precipitation and land use datasets. To assess structural uncertainty, we compared results of two phenology modules, LPJmL-OP and LPJmL-GSI, and demonstrated an improved ability of the latter for reproducing observed dynamics. Generally, different model structures from empirical models (Barichivich et al., 2014; Broich et al., 2014; e.g. de Jong et al., 2013a), model-data integration approaches (Caldararu et al., 2014; e.g. Stöckli et al., 2011) to process-oriented models (e.g. Lucht et al., 2002; Mao et al., 2013; Piao et al., 2006) were used to explain observed phenology and greenness dynamics. It has been shown that structural uncertainty is larger than parameter uncertainty in phenology modelling (Migliavacca et al., 2012). We reduced the uncertainty of FAPAR and phenology-related parameters by optimizing LPJmL-GSI against observations (Forkel et al., 2014). Consequently, modelled SOS and EOS were within the uncertainty of datasets in many regions (Figures 5). On the other hand, the model performance for peak FAPAR was usually lower than the data agreement. Peak FAPAR is related to FPC in LPJmL and thus depends on productivity, allocation, mortality, establishment, fire and land use change. These model routines were not yet improved through parameter optimization and thus provide potential for further model development. Other controls on greenness are not considered in LPJmL like nutrient availability (Fisher et al., 2012) or disturbances as insect infestations and storms (Bright et al., 2013; Eklundh et al., 2009), changes in grazing (Hilker et al., 2014) or in topography and soil conditions (Frost et al., 2014; Walker et al., 2009). However, we assume that these potential drivers, as for the case of fires, can be important regionally but are not of global importance for dynamics of land surface greenness and phenology. Increasing nutrient availability like increasing nitrogen deposition might be the reason why LPJmL underestimated peak FAPAR trends in comparison to the GIMMS3g in temperate forests of Europe (De Vries et al., 2006). Nevertheless, although nitrogen deposition changes globally

(Vitousek et al., 1997), nitrogen deposition was only identified of regional importance for trends in LAI in another modelling study (Mao et al., 2013). In summary, our approach considers uncertainties from different datasets and phenology methods, and allowed us for the first time to quantify the relative importance of light and water availability, observed fire activity, permafrost dynamics and LULCC on annual dynamics in land surface phenology and greenness.

#### **4.4.2 Importance of light, water availability, and LULCC for phenology dynamics**

Besides the known effect of seasonal cold temperature on SOS in temperate and boreal forests (Jeong et al., 2011; Keenan et al., 2014; Menzel et al., 2006; Wolkovich et al., 2012; Zhang et al., 2007), we demonstrated additionally the importance of seasonal light and water availability for inter-annual variability and trends in phenology. The availability and intensity of light affects spring leaf development especially in understory or late-successional species (Caffarra and Donnelly, 2011; Körner and Basler, 2010; Maeno and Hiura, 2000; Richardson et al., 2009). Moreover, our results suggest that the warming-induced advancing SOS trend is dampened by a limited availability of light in boreal forests. Thus, further warming in temperate and boreal regions might not necessarily result in a longer growing season (Körner and Basler, 2010; Richardson et al., 2013).

We found a strong effect of water availability and permafrost on SOS in boreal forests (Figure 4.8 i). The absence of permafrost in LPJmL increased water availability which resulted in stronger advancing SOS trends. Permafrost soils and the seasonal dynamic of the active layer are regulating the plant available water in many tundra and boreal forest regions and thus are affecting species composition and productivity (Benninghoff, 1952; Schuur et al., 2007; Sugimoto et al., 2002) and possibly phenology (Molau, 1997; Natali et al., 2012). Moreover, it has been shown that winter precipitation determines spring greenup in northern latitudes (Fu et al., 2014b). The equally important roles of seasonal light and water availability on trends and variance in SOS suggest that the temperature sensitivity of phenology might be overestimated if these factors are neglected in experimental studies (Wolkovich et al., 2012).

It is generally known that phenology depends on land cover type or species (Cleland et al., 2007; Körner and Basler, 2010; Richardson et al., 2013). However, the role of LULCC

on inter-annual dynamics of phenology was only little studied (Bradley and Mustard, 2008; Davison et al., 2011). Our findings emphasize, to our awareness for the first time, the control of LULCC on SOS in temperate regions (Figure 4.6 b). Therefore, we conclude that beside climatic drivers, anthropogenic factors need to be carefully considered to explain the inter-annual variability and trends of land surface phenology.

#### **4.4.3 Regional controls for greening and browning trends**

Seasonal water availability and LULCC had dominant effects on variance, correlation and trends in peak greenness in all regions (Figure 4.8). In boreal forests and tundra, permafrost and land cover change had the highest effects on variance and trends in peak FAPAR (Figure 4.8 a-c), despite temperature. Conversely to previous studies explaining greening trends in arctic and boreal regions with warming temperatures (Bhatt et al., 2013; Jia et al., 2009; Lucht et al., 2002; Xu et al., 2013), our results indicate that this warming is mediated into greening through cold temperature effects on phenology, and through changes in permafrost-regulated water availability and increasing vegetation coverage. The arctic and boreal greening is associated to an increase and intensification of shrub cover (Berner et al., 2013; Forbes et al., 2010; Myers-Smith et al., 2011). Warming increases the risk of permafrost degradation (Jorgenson et al., 2010; Shur and Jorgenson, 2007) and the associated increases in water availability in lowlands and nutrient availability contribute to greening trends (Berner et al., 2013; Frost et al., 2014; Reynolds et al., 2013; Walker et al., 2009). The importance of these effects could be a potential explanation for the weakening relationship between temperature and greenness dynamics (Piao et al., 2014). The influence of permafrost dynamics on greening trends contributes to the divergent continental patterns of less greening/more browning in boreal forests of North America than in Eurasia (Bi et al., 2013) because boreal forests of Eurasia are too larger extent underlain by permafrost.

In temperate forest of Europe, inter-annual variability and trends in peak FAPAR were mostly related to LULCC (Figure 4.8 d). Indeed, large areas experienced forest regrowth in the last decades (Fuchs et al., 2013). In central and eastern Europe, this was induced through socio-economic transitions in the former socialist states from the 1980

to the 1990s: croplands were reforested (Alcantara et al., 2013; Kozak, 2003; Kuemmerle et al., 2008) and forests experienced a recovery after damages by air pollution in the 1980s (Kubíková, 1991; Main-Knorn et al., 2009; Polák et al., 2006; Schulze, 1989). Additionally, disturbance-induced land cover changes like storms and insect infestations damaged especially the dominant spruce forests (Jönsson et al., 2009; Wermelinger, 2004). This effect can be also seen in satellite-derived vegetation indices (Eklundh et al., 2009). Although not all of these disturbances that affect land cover are represented in LPJmL, our results demonstrate that fixed land cover conditions of 1982 cannot explain the observed greening trends in temperate forests of Europe.

In the Sahel, peak FAPAR was equally affected by seasonal water availability and LULCC (Figure 4.8 e). This is in agreement with previous findings of re-greening in 1982-2011 because of increasing precipitation trends after a period of severe droughts (Eklundh and Olsson, 2003; Herrmann et al., 2005; Hickler et al., 2005; Olsson et al., 2005). Opposing to this precipitation-induced greening trends, increased land degradation occurs (Dardel et al., 2014a). Increasing trends in precipitation and land degradation are accepted as the main controls for the variability and trends in greenness in the Sahel, but the relative contribution of these opposing factors is debated (Brandt et al., 2014a; Dardel et al., 2014a; Fensholt et al., 2013). Some studies associate the greening to a recovery of trees (Brandt et al., 2014b), but it has been shown that changes in tree cover depend highly on the used satellite dataset with diverging results (Horion et al., 2014). Our results demonstrate that changes in vegetated area are more important for the trend in peak FAPAR than seasonal changes in water availability.

In both tropical forest regions (Amazon and Congo), seasonal water availability and LULCC were the dominant controls for variance and trends in peak FAPAR (Figure 4.8 f-g). Although some datasets have significant trends in vegetation greenness over the Amazon and the Congo basin (Figure 4.3), the satellite datasets have a poor agreement in temporal dynamics. Greenness observations over tropical forest regions are affected from several observational limitations (Samanta et al., 2012) and seasonal changes have been identified as artefacts from sun-sensor geometries (Morton et al., 2014). Consequently, previous hypotheses about a greening of the Amazon during drought periods because of increased light availability (Huete et al., 2006; Myneni et al., 2007;



Saleska et al., 2007) have been falsified (Morton et al., 2014). Moreover decreased productivity and increased tree mortality was reported during drought periods (Nepstad et al., 2007; Phillips et al., 2009; Zhou et al., 2014). Our results support these findings that water availability rather than light affect seasonal to decadal variations in peak greenness in tropical forests.

#### **4.4.4 Controls for global greening trends and consequences for prognostic modelling**

Our results demonstrate that global greening and phenology trends are mostly associated to 1) annual changes in land cover and land use, and 2) to seasonal effects of cold temperature and water availability on phenology (Figure 4.9 e-f). This finding partly agrees with previous studies about warming-induced greening trends in high-latitudes conducted at regional scale (Lucht et al., 2002; Xu et al., 2013). However, we additionally emphasize the strong contributions of LULCC and water availability on inter-annual and decadal variability of vegetation greenness globally. Thus, greening trends are linked to an intensification of the global water cycle (Huntington, 2006) that leads to regional increasing precipitation (New et al., 2001), soil moisture (Sheffield and Wood, 2008) or earlier soil thaw dates (Smith et al., 2004). Land use changes have global consequences for ecosystem structure, functions and services (Foley et al., 2005). Important changes in forests cover have been reported for the last decade (Hansen et al., 2013) that should clearly be represented in the inter-annual variability of land surface phenology and greenness. Additionally, changes in anthropogenic land use affected the inter-annual variability and trends in land surface phenology and greenness especially in agricultural regions of Europe, the US, south-east Asia, sub-Saharan Africa and South America (Mueller et al., 2014; Wang et al., 2014). Although CO<sub>2</sub> fertilization clearly affected the trend in global greenness in 1982-2011, the effect of CO<sub>2</sub> fertilization was small in comparison to seasonal temperature and water effects, land use and land cover change.

In conclusion, our study reveals spatially and temporally distinct drivers of land surface phenology and greenness dynamics. The importance of a driver varies according to temporal scale, region, and metric of interest (start of season, end of season, peak greenness). Thus future studies should clearly state the investigated

aspect of phenology. In temperate and boreal forests, seasonal light and water availability, in addition to seasonal cold temperature, mainly control inter-annual variability and trends of the start of growing season. In all regions, inter-annual variability and trends of peak greenness are driven by seasonal water limitation and land use and land cover change (LULCC). The large importance of LULCC and water availability on phenology and greenness dynamics requires a better observation-based quantification and understanding of land transitions, natural vegetation dynamics, and water-vegetation couplings in order to improve prognostic land surface models. Current prognostic land surface models have large uncertainties about the future development of the terrestrial carbon cycle (Friedlingstein et al., 2006, 2014). These uncertainties are partly related to an underestimation of the role of water in comparison to temperature in ecosystem carbon turnover (Carvalhais et al., 2014) and to different representations of vegetation transitions (Friend et al., 2014). Our findings emphasize that in addition to limitation by low temperatures, water availability is globally a co-dominant control for start of growing season and greening trends. These results point towards the reformulation of phenology models contributing to a better prognostic description of vegetation controls on the global carbon cycle. The model introduced and tested here overcomes some of these limitations and missing drivers.

## **Acknowledgments**

We are grateful to Ranga Myneni and colleagues for providing the GIMMS3g FAPAR dataset. We thank Trevor F. Keenan for his comments on phenology methods. We thank Werner von Bloh for help in setting up LPJmL model simulations. MF was funded from the EU FP7 Carbones project (242316) and conducted this work under the International Max Planck Research School for Global Biogeochemical Cycles.



# **5 Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere**

---

This chapter is submitted as:

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R. F., Heimann, M., Thonicke, K., Zaehle, S. and Reichstein, M.: Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere, submitted to Nature, 2015-01-30.

Methods and Extended Data of this chapter can be found in chapter 9.

## 5      **Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere**

### **Abstract**

The seasonality of atmospheric carbon dioxide (CO<sub>2</sub>) in the northern hemisphere is mainly controlled by photosynthetic carbon uptake and respiratory release of the terrestrial biosphere (Keeling et al., 1996). Monitoring of atmospheric CO<sub>2</sub> concentrations has shown an amplification of the seasonal CO<sub>2</sub> cycle in high northern latitudes (> 40°N) since the 1960s (Bacastow et al., 1985; Graven et al., 2013; Keeling et al., 1996). However, the mechanisms behind this amplification of the CO<sub>2</sub> amplitude are not yet fully understood (Graven et al., 2013; Randerson et al., 1997). Here we show by coupling a terrestrial biosphere model (LPJmL) to an atmospheric transport model (TM3) that the larger CO<sub>2</sub> amplitude trend in high northern latitudes compared to lower latitudes is mainly driven by an enhancement of photosynthetic carbon uptake in boreal and arctic regions. The interaction of direct climate effects on photosynthesis and indirect effects on ecosystem productivity through changes in vegetation coverage are able to explain the observed trend, which results in increasing net biosphere carbon uptake during the boreal summer. CO<sub>2</sub> fertilization and agricultural land-use change (Gray et al., 2014; Zeng et al., 2014) contribute moderately to the CO<sub>2</sub> amplitude trends, but these factors cannot explain the latitudinal gradient of the trend. Our results emphasize the importance of the high-latitude climate-vegetation-carbon cycle feedbacks, and indicate that during the last decades photosynthetic carbon uptake has more reacted strongly to high-latitude warming than carbon release processes implying a sustained negative feedback.

### **Main text**

Atmospheric monitoring has shown an increase in the atmospheric abundance of CO<sub>2</sub> due to anthropogenic emissions but also an increase in the amplitude of the seasonal cycle of CO<sub>2</sub> in the northern hemisphere (IPCC, 2014; Keeling et al., 1996). Based on

evidence from surface stations as well as integrative flight campaigns in the middle and upper troposphere (Graven et al., 2013; Keeling et al., 1996), the observations consistently show that the amplitude of the seasonal cycle of CO<sub>2</sub> (hereinafter shortly called “CO<sub>2</sub> amplitude”) has increased more strongly in high latitudes (50% increase north of 45°N) (Graven et al., 2013) than in low latitudes (< 25% increase south of 45°N) (Graven et al., 2013) since the 1960s. The CO<sub>2</sub> amplitude in the northern hemisphere is primarily driven by the difference in phase and amplitude between photosynthetic carbon uptake (gross primary production, GPP) and ecosystem respiration (Reco) by the land biosphere (Heimann et al., 1989; Randerson et al., 1997). The strong seasonality of these processes in northern land ecosystems causes the average CO<sub>2</sub> amplitude to increase from low to high northern latitudes (Randerson et al., 1997).

Changes in the CO<sub>2</sub> amplitude are thus likely an indicator for trends in the terrestrial carbon cycle of northern ecosystems. In particular, boreal and arctic regions experienced strong warming trends in the last decades (IPCC, 2014) (Figure 1a). In the same period an enhanced plant growth was observed from satellites (Lucht et al., 2002; Myneni et al., 1997; Xu et al., 2013) which is usually referred as “greening” (Figure 1b). Satellite observations of a greening land surface are confirmed by ground observations of an increasing coverage of deciduous shrubs in the tundra and along the arctic/boreal tree line (Berner et al., 2011; Myers-Smith et al., 2011). Accordingly to the increase in vegetation cover, independent estimates of GPP from machine learning approaches (Jung et al., 2011) and of net biome productivity (NBP) from atmospheric inversions (Rödenbeck et al., 2003), respectively, demonstrate an increasing terrestrial gross and net carbon uptake (Figure 1c-d). These multiple observational signals demonstrate an amplification of plant productivity in northern ecosystems which could likely cause the increase in CO<sub>2</sub> amplitude (Bacastow et al., 1985; Barichivich et al., 2013; Keeling et al., 1996; Myneni et al., 1997). A recently proposed alternative hypothesis is that the intensification of agriculture in mid-latitudes contributes to CO<sub>2</sub> amplitude trends (Gray et al., 2014; Zeng et al., 2014). Accordingly to these studies, the intensification of agriculture explains around 20% (Gray et al., 2014) or up to 45% (Zeng et al., 2014) of the observed mean global amplitude trend, depending on methodology used. However, these do not comprehensively account for the importance of other factors like changes in vegetation cover in high latitude ecosystems, fossil fuel emissions or

atmospheric transport. Other plausible contributing factors are the direct effects of climate change and CO<sub>2</sub> fertilization on plant photosynthesis and ecosystem respiration, as well as an indirect effect of increasing vegetation coverage on ecosystem productivity. Current earth-system models consistently underestimate the amplitude trend (Graven et al., 2013), and a comprehensive explanation of the patterns of CO<sub>2</sub> amplitude changes based on these factors is still missing.

Here we examine this question by combining observations from long-term monitoring stations of atmospheric CO<sub>2</sub> concentration, satellite observation of vegetation greenness (Zhu et al., 2013), global observation-based datasets of GPP (Jung et al., 2011), and NBP (Rödenbeck et al., 2003) with the LPJmL dynamic global vegetation model (Bondeau et al., 2007; Sitch et al., 2003) and the TM3 atmospheric transport model (Heimann et al., 1989) (hereinafter called LPJmL+TM3, see Methods) to explain the latitudinal gradient of CO<sub>2</sub> amplitude trends. Unlike other biosphere models that were previously evaluated against CO<sub>2</sub> amplitude trends (Graven et al., 2013; Zeng et al., 2014), LPJmL considers several processes that potentially contribute to a better explanation of these trends like agriculture, irrigation and land use change (Bondeau et al., 2007); vegetation dynamics; and important processes for boreal and arctic ecosystems like permafrost (Schaphoff et al., 2013), fire and post-fire succession (Thonicke et al., 2010). Moreover, LPJmL uses an improved phenology module (Forkel et al., 2014), which has been optimized against satellite observations of FAPAR (fraction of absorbed photosynthetic active radiation), albedo, and an observation-based data set of GPP, altogether resulting in a better representation of vegetation greenness dynamics, tree cover, as well as global carbon fluxes and stocks (Forkel et al., 2014). Note that atmospheric CO<sub>2</sub> data were not used to constrain LPJmL.

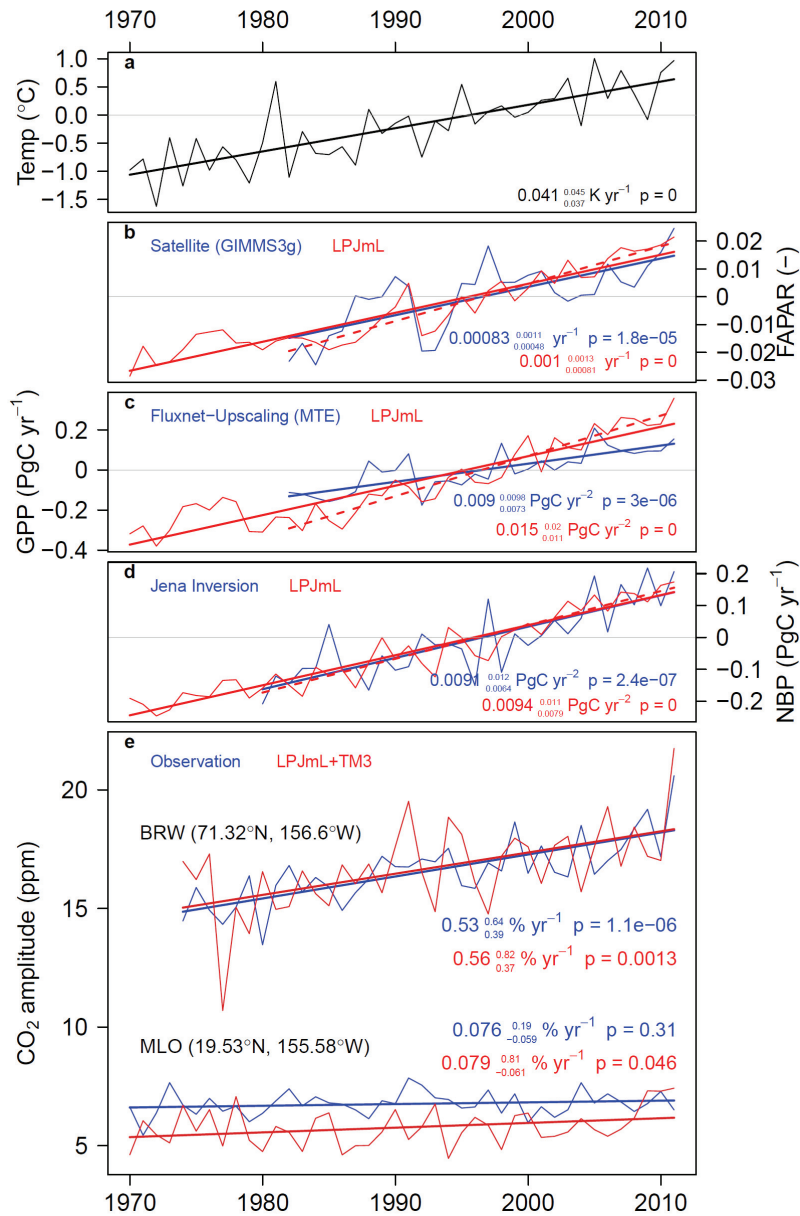


Figure 5.1: Amplification of plant activity in the northern biosphere.

Annual time series and linear trends of (a) mean annual air temperature, (b) peak FAPAR (fraction of absorbed photosynthetic active radiation), (c) annual amplitude of gross primary production (GPP), (d) annual amplitude of net biome productivity (NBP), and (e) seasonal amplitude of atmospheric CO<sub>2</sub> at Barrow (BRW) and Mauna Loa (MLO). Time series in (a-d) were spatially aggregated for boreal and arctic land regions north of 41°N, and the 1982-2011 mean has been subtracted. Dashed lines indicate the trends for the overlapping period of LPJmL simulations and observations. P-values were calculated with the Mann-Kendall trend test.

We estimated CO<sub>2</sub> amplitude trends in observed time series at 19 monitoring stations with at least 20 years of data (Extended Data Table 1, Methods on CO<sub>2</sub> time series processing and site selection). We found much stronger positive CO<sub>2</sub> amplitude trends at high latitude stations (e.g. 0.093 ppm yr<sup>-1</sup> ≈ 0.53% yr<sup>-1</sup> at Point Barrow, BRW in 1971-2011) than at low latitude stations (e.g. 0.072 ppm yr<sup>-1</sup> ≈ 0.076% yr<sup>-1</sup> at Mauna Loa,



MLO, in 1970-2011, Figure 1e). These estimated trends were of similar magnitude to previous studies (Graven et al., 2013; Randerson et al., 1997) although not identical because of differences in the used station records, time series analysis methods, and especially time series length. Especially at MLO we found much weaker CO<sub>2</sub> amplitude trends because this trend originates mostly from low CO<sub>2</sub> amplitude values in the 1960s and disappears from 1970 onwards. To account for the effect of time series length, we estimated the uncertainties in CO<sub>2</sub> amplitude trends by computing trends for different start and end years (Methods on trend analysis, Extended Data Figure 1). The estimated uncertainties ( $0.53^{0.64}_{0.39}$  % yr<sup>-1</sup> at BRW,  $0.076^{0.19}_{-0.059}$  % yr<sup>-1</sup> at MLO, 50<sub>2.5</sub><sup>97.5</sup> %-iles of trend slope ensemble) demonstrate that only high latitude stations have persistent long-term increases in CO<sub>2</sub> amplitude. We found a large longitudinal variability with partly large uncertainties in CO<sub>2</sub> amplitude trends at mid-latitude stations (30-60°N) (Figure 2a) which was already shown previously based on much shorter time series (Randerson et al., 1997).

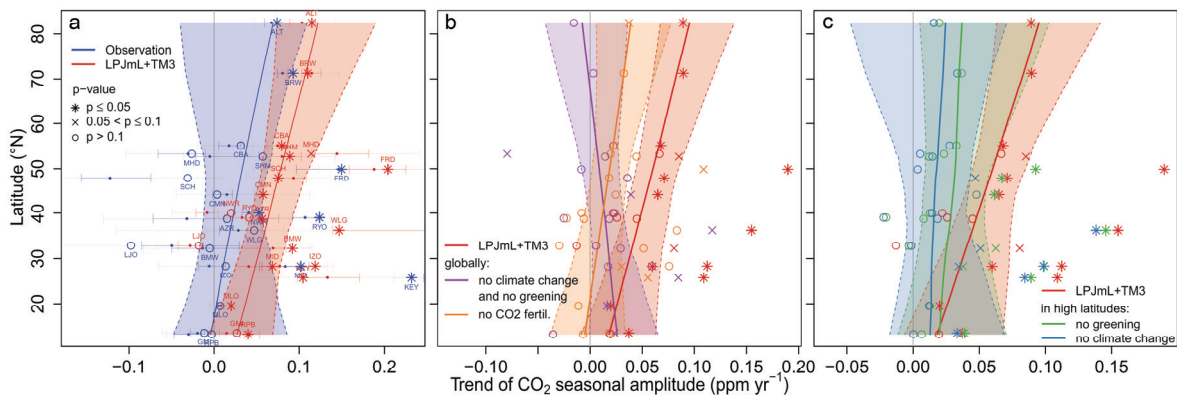


Figure 5.2: Latitudinal gradients of trends in the seasonal CO<sub>2</sub> amplitude and its drivers.

- (a) Simulated and observed CO<sub>2</sub> amplitude trends with 95% confidence intervals. (b) Global effects of CO<sub>2</sub> fertilization and climate change on the latitudinal gradient. Removing the effect of CO<sub>2</sub> fertilization on photosynthesis reduces CO<sub>2</sub> amplitude trends globally but has no effect on the latitudinal gradient. The latitudinal gradient disappears with a constant climate. (c) Separation of the effects of greening and climate change in high-latitude regions on the latitudinal gradient. The latitudinal gradient disappears both without high-latitude greening (i.e. constant vegetated area but still climate change) and without climate change (i.e. no climate change but still forcing changes in vegetated area).

In comparison to the observations, LPJmL+TM3 had on average stronger CO<sub>2</sub> amplitude trends than the observations at station level (Figure 2a). It nevertheless

reproduced the observed pattern of strong positive CO<sub>2</sub> amplitude trends north of 45°N, the large variability of trends in mid-latitudes, and small or no trends south of 20°N. LPJmL+TM3 had also similar CO<sub>2</sub> amplitude trends at higher atmosphere levels like previously observed (Graven et al., 2013) (Extended Data Figure 2). We found that simulations of CO<sub>2</sub> amplitudes were sensitive to the choice of the meteorological forcing dataset for the TM3 transport model (Extended Data Figure 3a). Therefore we propagated the uncertainty both from time series length and meteorological forcing to the overall uncertainty of simulated CO<sub>2</sub> amplitude trends for a more robust model evaluation. We additionally checked if the inter-annual variability of ocean CO<sub>2</sub> uptake could contribute to CO<sub>2</sub> amplitude trends but found no distinct contribution in comparison to just using a climatology of ocean uptake (Extended Data Figure 3b). LPJmL+TM3 simulations were well correlated with observations regarding spatial patterns of mean CO<sub>2</sub> amplitude values ( $r = 0.84$ ) and trends ( $r = 0.51$ ) (Extended Data Figure 4). LPJmL also reproduces observations of the amplification of plant productivity in northern ecosystems like the observed greening trend (Forkel et al., 2014). LPJmL simulates increasing coverage of deciduous vegetation in arctic and boreal regions as seen in ground observations, and also agrees with independent observation-based estimates of trends in the amplitude of GPP and NBP (Figure 1b-e, Extended Data Figure 5). We conclude from these comparisons of model simulations against multiple observational data sets that LPJmL-TM3 can be applied to comprehensively assess controlling factors for CO<sub>2</sub> amplitude trends.

We quantified the contribution of terrestrial NBP from different regions to CO<sub>2</sub> amplitude trends (Extended Data Table 3a, Extended Data Figure 6b, Methods on Factorial model experiments and quantification of contributions). Boreal NBP contributed 51% to the average CO<sub>2</sub> amplitude in the 1970s at northern high latitude sites (> 45°N). In the 2000s, the contribution of boreal regions increased to 54% which is related to an annual increase of  $0.35^{0.41}_{0.14} \% \text{ yr}^{-1}$ . The contribution of arctic NBP was 17% in the 1970s and increased with  $0.18^{0.24}_{0.12} \% \text{ yr}^{-1}$ . NBP from global agricultural regions contributed with 11% in the 1970s and increased with  $0.14^{0.16}_{0.01} \% \text{ yr}^{-1}$ . Temperate and tropical regions had only minor contributions to the trends in CO<sub>2</sub> amplitude. However, only trends in the contribution of arctic and boreal NBP were significant ( $p =$

0.03 and  $p = 0.01$ , respectively) at northern sites. At low latitude sites ( $0 - 45^{\circ}\text{N}$ ), NBP from boreal regions still contributed dominantly to the increase in  $\text{CO}_2$  amplitude with  $0.46_{0.09}^{0.52} \% \text{ yr}^{-1}$  ( $p = 0.12$ ), followed by NBP from agricultural regions with  $0.17_{-0.21}^{0.26} \% \text{ yr}^{-1}$  ( $p = 0.37$ ) and NBP from arctic regions with  $0.14_{0.08}^{0.19} \% \text{ yr}^{-1}$  ( $p = 0.12$ ). Therefore boreal regions contributed approximately with 57%, arctic regions with 25% and agricultural regions with 17% to the overall  $\text{CO}_2$  amplitude trend at northern latitude sites (41%, 14% and 20% at low latitude sites, respectively). These contributions from agricultural regions provide support for the lowest of the previous estimates (17%-25%) (Gray et al., 2014) for the contribution to the increase in  $\text{CO}_2$  amplitude from increased cropland productivity. We found no profound change in the contribution from fossil fuel emissions or ocean  $\text{CO}_2$  uptake to the trends in  $\text{CO}_2$  amplitude. These results demonstrate that boreal and arctic regions dominate the increase in the  $\text{CO}_2$  amplitude globally. Consequently, besides agriculture, processes such as permafrost dynamics (Schaphoff et al., 2013) or phenology (Forkel et al., 2014), that are of major importance for vegetation productivity in boreal and arctic regions, strongly affected the simulation of  $\text{CO}_2$  amplitude trends (Extended Data Figure 6c), suggesting a strong climate control on long term ecosystem dynamics.

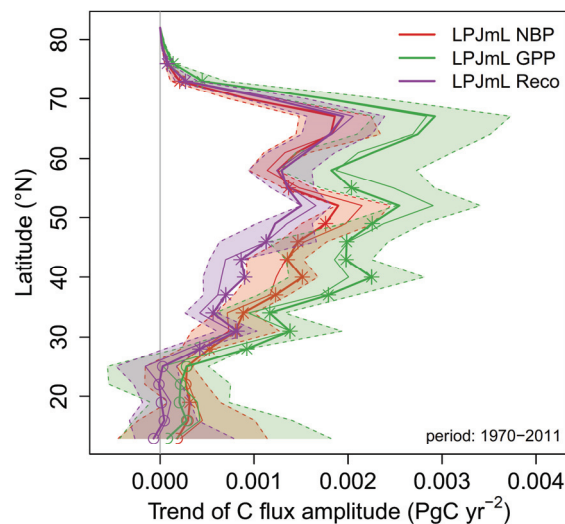


Figure 5.3: Trends in the seasonal amplitude of terrestrial net biome productivity (NBP), gross primary productivity (GPP) and ecosystem respiration (Reco) per latitudinal bands of  $3^{\circ}$  width. Uncertainty bands represent 95% confidence intervals of the amplitude trend depending on time series length.

Both GPP and ecosystem respiration (Reco) can potentially contribute to the increasing CO<sub>2</sub> amplitude. We found stronger GPP than Reco amplitude trends across all northern latitudes (Figure 3) and therefore GPP had a stronger contribution than Reco to the CO<sub>2</sub> amplitude (Extended Data Table 3b). For example, GPP contributed with 250% and Reco with -154% to the CO<sub>2</sub> amplitude across northern high latitude sites in the 1970s. The sum of these two contributions shows that terrestrial net ecosystem exchange (NEE) contributed with 96% to the CO<sub>2</sub> amplitude. In the 2000s, terrestrial NEE contributed with 97% (= 234% - 137%, GPP - Reco) to the seasonal CO<sub>2</sub> amplitude at northern latitudes. The increasing terrestrial net contribution is caused by an increase in the contribution of GPP with a trend of  $2.3_{-0.1}^{3.2}$  % yr<sup>-1</sup> whereas the contribution of Reco showed a weaker increase ( $-1.5_{-2.24}^{0.22}$  % yr<sup>-1</sup>, a negative sign for a trend in the contribution of Reco to the CO<sub>2</sub> amplitude indicates an increase because of the negative overall contribution of Reco to the seasonal cycle). The increasing contribution of the net land uptake on the CO<sub>2</sub> amplitude is reflected by an increase of total net biome productivity (NBP) in northern lands (> 41°N) of  $0.011_{0.002}^{0.021}$  PgC yr<sup>-1</sup> (LPJmL) and  $0.016_{0.011}^{0.024}$  PgC yr<sup>-1</sup> (Jena Inversion) per year. These results suggest that the positive trends in the seasonal CO<sub>2</sub> amplitude can be explained by a photosynthesis-driven increase in net uptake of the northern biosphere.

Several factors can contribute to the increased photosynthetic carbon uptake, and thus the latitudinal gradient of the increasing CO<sub>2</sub> amplitude. Rising atmospheric CO<sub>2</sub> and warming that relax climatic limits for plant growth in northern ecosystems are direct physiological effects that can enhance photosynthesis (Norby et al., 2005). To test the relative effect of CO<sub>2</sub> fertilization and climate change on CO<sub>2</sub> amplitude trends, we performed two model experiments with LPJmL where we 1) kept temperature and precipitation at 1965-1975 levels for the period 1970-2011 (i.e. no climate change); and 2) held CO<sub>2</sub> constant at 325.7 ppm after 1970 to quantify the effect of CO<sub>2</sub> fertilization (Figure 2b). We found that both climate change and CO<sub>2</sub> fertilization affect CO<sub>2</sub> amplitude trends, but with regional differences: Climate change was the dominant factor on CO<sub>2</sub> amplitude trends north of 40°N. For example, the simulated trend increased from  $-0.05_{-0.27}^{0.18}$  % yr<sup>-1</sup> without climate change to  $0.56_{0.37}^{0.82}$  % yr<sup>-1</sup> with climate change at Barrow. On the hand, CO<sub>2</sub> fertilization was the dominant factor south of

40°N. For example, the simulated trend increased from  $-0.5_{-0.67}^{0.22}$  % yr<sup>-1</sup> without CO<sub>2</sub> fertilization to  $0.079_{-0.061}^{0.81}$  % yr<sup>-1</sup> with CO<sub>2</sub> fertilization at Mauna Loa. Without the effect of CO<sub>2</sub> fertilization, CO<sub>2</sub> amplitude trends were generally lower (ca. -0.04 ppm yr<sup>-1</sup> across all latitudes) but the latitudinal gradient of stronger CO<sub>2</sub> amplitude trends in northern than in southern regions was not affected. Without the effect of climate change, the strong CO<sub>2</sub> amplitude trends in northern latitudes disappeared and thus the latitudinal gradient reverted towards stronger trends south of 40°N (Figure 2b). These results show that the much stronger CO<sub>2</sub> amplitude trends at northern latitudes are mainly dominated by NBP from boreal and arctic regions, from gross primary production and climate change. We further quantified the role of an increasing vegetation coverage (or “greening”) in boreal and arctic regions on CO<sub>2</sub> amplitude trends. To disentangle the effect of climate change and greening in boreal and arctic regions, we performed two more model experiments: 1) We again fixed climate in the period 1970-2011 according to the temperature and precipitation conditions in 1965-1975 but prescribed changes in vegetated area as simulated in the standard model run (i.e. no climate change but still greening). 2) We used observed climate but fixed vegetated area after 1970 (i.e. no greening but still climate change). Both experiments were performed only for boreal and arctic regions while the rest of the world was simulated following the normal simulation protocol. The latitudinal gradient of stronger CO<sub>2</sub> amplitude trends in northern latitudes disappeared without the effects of greening and climate change in northern ecosystems, respectively (Figure 2c). The effect of climate change in northern regions had a higher importance than the greening effect on photosynthesis. However, this result demonstrates that the increase in vegetated area was a major contribution especially to the strong CO<sub>2</sub> amplitude trends in northern latitudes. This relationship between greening in northern ecosystems and CO<sub>2</sub> amplitude trends has been suggested previously based on observations (Barichivich et al., 2013; Myneni et al., 1997) but the attribution to regional drivers and mechanisms has been not treated comprehensively before.

Our results demonstrate that the latitudinal gradient of CO<sub>2</sub> amplitude trends is ultimately shaped by the interaction of recent climate change and vegetation dynamics in boreal and arctic ecosystems. Overall, we offer a parsimonious and plausible explanation of the decadal trend of seasonal CO<sub>2</sub> amplitude in the high latitudes based

on a comprehensive consideration of several possible factors. Our analysis indicates that in the last decades photosynthetic carbon uptake reacted more strongly to changes in climate than carbon release processes implying a sustained negative feedback. Stimulation of photosynthesis and greening by warming and associated drivers can however not be unlimited, because of limitations by photoperiod, nutrients and possibly increased mortality. Thus at some point in the future the positive trend in the seasonal CO<sub>2</sub> amplitude might stall. Continued long-term observation of atmospheric CO<sub>2</sub> together with satellite observations of productivity and vegetation dynamics will be the key to detect such change in the high-latitude carbon cycle dynamics.

### **Acknowledgements**

We thank the following institutions and scientists that acquired and provided CO<sub>2</sub> station data: Ed Dlugokencky (NOAA), Scripps Institution of Oceanography, Doug Worthy (Environment Canada), Frank Meinhardt (Umweltbundesamt), Ray Langenfelds and Paul Krummel (CSIRO), and Hiroshi Koide (JMA). We further thank the GIMMS group, European Commission/Joint Research Centre/Netherlands Environmental Assessment Agency, and Martin Jung (MPI-BGC) for sharing their datasets. We thank Sibyll Schaphoff (PIK) for comments on LPJmL. We acknowledge financial support by the European Union (FP7) through the projects GEOCARBON (283080), CARBONES (242316) and EMBRACE (283201).



## 6 Conclusions and future research

---

### **How do observational uncertainties contribute to the quantification of dynamics in land surface phenology and greenness?**

*The consideration and quantification of observational uncertainties in the analysis of dynamics in land surface phenology and greenness are essential to ensure the reliability of scientific results.*

### **How can DGVMs be improved to better represent phenology and greenness dynamics?**

*The LPJmL DGVM required the implementation and optimization of a new phenology model to better represent greenness dynamics on seasonal to decadal time scales. Thereby it was important to consider water availability as an additional control on phenology. In addition, these model improvements also resulted in a better representation of global carbon cycle observations.*

### **What are the controls on land surface phenology and greenness on multiple time scales?**

*The importance of environmental controls for land surface phenology and greenness varies regionally and depends on time scale. However, water availability and temperature play co-dominant roles globally for seasonal dynamics, inter-annual variability and trends in land surface phenology and greenness.*

### **What is the contribution of high latitude greening on the increasing amplitude of atmospheric CO<sub>2</sub>?**

*The increasing seasonal amplitude of atmospheric CO<sub>2</sub> in northern latitudes is mainly driven by enhanced terrestrial gross primary production in boreal and arctic ecosystems, which is caused by the climate-induced greening trend.*



## 6 Conclusions and future research

### 6.1 Accounting for observational uncertainties

*The consideration and quantification of observational uncertainties in the analysis of dynamics in land surface phenology and greenness are essential to ensure the reliability of scientific results.*

Differences in satellite datasets affect estimates of land surface phenology and greenness on all temporal scales from instantaneous to mean values, seasonal cycles, inter-annual variability and trends (Figures 4.2, 4.3; Tables 4.2, 4.3). The use of a single satellite dataset can therefore result in biased conclusions about ecosystem dynamics and model performances when applied in the evaluation of ecosystem models.

The availability and varying performance of time series analysis methods is a major source of observational uncertainty that needs to be quantified. The results demonstrate that the performance of trend change detection methods in detecting the real underlying trend and breakpoints decreases as the inter-annual variability of the time series increases (Figure 2.7, Table 2.2). Specifically, all methods were prone to detect trends opposite to the underlying reality (e.g. detecting browning instead of greening) (Table 2.1). Nevertheless, detected breakpoints in NDVI trends in Alaska coincide with drought periods, fire events, but also with sensor changes in the NDVI record, and the amount of flagged NDVI values due to low data quality (Figures 2.9 and 2.10; chapter 2.4.4). Thereby the use or exclusions of snow-flagged NDVI values had a strong effect on the estimation of trends and breakpoints (Figure 2.8). Given the strong dependence of trend-detection method performance on inter-annual variability, a different trend analysis approach was used in chapter 5 and is proposed for future studies (Figure 9.1). This approach computes an ensemble of trend slopes for different combinations of years of the time series and thus quantifies the uncertainty of a trend by accounting both for the effects of time series length and inter-annual variability. In summary, given the low reliability of trend change detection methods, satellite-derived vegetation greenness time series need to be screened for low quality values (snow, clouds) and the uncertainty of estimated trends needs to be quantified by either using

ensembles of trend methods or by directly accounting for inter-annual variability and time series length.

The varying performance of phenology detection methods introduces major uncertainties in estimates of phenology events like the start or end of growing season and their temporal dynamics. Based on previous findings about the reliability of phenology methods for vegetation index time series (e.g. de Beurs and Henebry, 2010; Kandasamy et al., 2013; White et al., 2014, 2009), an ensemble of phenology methods and satellite datasets was used in this study (chapter 4.3.1). Enormous differences in the timing of the start or end of growing season were found by using different phenology methods and datasets (chapter 4.3.1). As it is almost impossible to identify a single best phenology method for continental or global-scale applications, the use of an ensemble of several satellite datasets and phenology methods is a robust approach to assess dynamics in land surface phenology and greenness and to evaluate ecosystem models.

Previous satellite-based studies of greening and phenology trends rarely accounted for observational uncertainties, i.e. they just relied on one satellite dataset and analysis method (e.g. Huete et al., 2006; de Jong et al., 2011, 2013; Piao et al., 2011; Wang et al., 2011). By considering differences between satellite datasets some of these studies need to be critically re-evaluated or have been refuted (Guay et al., 2014; Jiang et al., 2013; Morton et al., 2014). For example, strong browning trends in boreal forests in the GIMMS dataset are not confirmed by the newer GIMMS3g dataset or by datasets from other sensors (Guay et al., 2014; Jiang et al., 2013) (Table 4.3, Figure 4.3). Additionally, previous observations of a seasonal green-up of the Amazon under drought conditions and increased shortwave radiation (Huete et al., 2006; Myneni et al., 2007) have been rebutted (Morton et al., 2014; Samanta et al., 2010) and are neither supported by the results of this thesis (chapter 4.4.3). In summary, in future studies about land surface phenology and greenness dynamics, it is imperative that observational uncertainties from different satellite datasets and time series analysis methods are always assessed and evaluated against alternative observational datasets in order to avoid wrong or biased conclusions.

Several alternative data sets that are valuable to evaluate findings from coarse-scale satellite greenness time series and to get a better understanding of phenology and

vegetation dynamics already exist or will become available in the near future. As this thesis shows, vegetation dynamics or land cover change (1) strongly affect land surface phenology and greenness, and (2) interacts with climate dynamics, resulting in positive trends of the terrestrial net carbon uptake in high-latitude ecosystems. However, the current availability of datasets impedes the evaluation of such model simulations of vegetation-carbon cycle dynamics. Current land cover maps usually have no temporal domain and hold several uncertainties that prohibit their use to assess land cover change (Congalton et al., 2014). On the other hand, re-analyses of land cover change often depend on uncertain country-level statistics, have coarse spatial resolution, or are only available at high resolution only regionally (e.g. Fuchs et al., 2014 for an overview of available datasets). Alternatively, the long time series of Landsat sensors since 1972 could be used but is currently under-exploited for large scale analyses of vegetation dynamics. Landsat images were used to assess long-term (30 years) greenness trends regionally (Fraser et al., 2011, 2014; Verbyla, 2008), to assess regional changes in vegetation cover (1973-2012) (Urban et al., 2014), and to quantify changes in tree cover globally but for a shorter time period only (2000-2012) (Hansen et al., 2013). The higher spatial resolution of Landsat data provided a better understanding of the regional variability of greening and their dependence on vegetation types (Fraser et al., 2014; Urban et al., 2014). Consequently, future efforts within remote sensing should aim to harmonize historical high resolution satellite images with recent (Landsat 8, Proba-V) and future high resolution multi-spectral satellite data (e. g. from Sentinel-2, Drusch et al. (2012)). Such harmonization efforts can provide consistent half-century time series of vegetation indices and land cover change globally. The higher spatial resolution of these datasets might provide a better insight in the spatial dependencies of vegetation dynamics and will be ultimately the basis to account for land cover change in re-analyses of terrestrial carbon fluxes regardless of the used modelling approach.

Another alternative data stream for future analyses of land surface phenology (i.e. for the evaluation of satellite remote sensing data or for model-data integration) is obtained from the PhenoCam network (Klosterman et al., 2014; Migliavacca et al., 2011; Sonnentag et al., 2012). PhenoCam is a network of measurement sites at which photographic images of an ecosystem (“near-surface” remote sensing) are taken automatically several times per day. High-temporal resolution time series of canopy

greenness can be derived from these images. Although PhenoCam cannot provide the large spatial coverage of satellite remote sensing, it provides phenological observations with less uncertainty because of the higher temporal resolution, the reduced effect of atmospheric distortions on observations, and the higher spatial details that allow the distinction of vegetation patches or individual plants. As PhenoCam time series become longer with ongoing measurements, they will become valuable for the assessment of inter-annual variability and trends of ecosystem or species phenology.

## **6.2 From phenology modelling to multi-scale model-data integration**

*The LPJmL DGVM required the implementation and optimization of a new phenology model to better represent greenness dynamics on seasonal to decadal time scales. Thereby it was important to consider water availability as an additional control on phenology. In addition, these model improvements also resulted in a better representation of global carbon cycle observations.*

The original phenology model of the LPJmL DGVM (LPJmL-OP) was unable to reproduce observed greenness dynamics especially in grasslands (Figure 3.4). It has been shown that water availability needs to be considered as a controlling factor for the phenology of perennial dry ecosystems or grasslands (Liu et al., 2013; Richardson et al., 2013; Yuan et al., 2007). Therefore we adapted the GSI phenology model (Jolly et al., 2005) for LPJmL (LPJmL-GSI) but introduced a new functional relationship that depends on water availability (chapter 3.2.2.2). As this is an empirical approach to simulate phenology, model parameters need to be estimated. Parameters that account for phenology, canopy light absorption, albedo, and gross primary productivity were estimated to improve model simulations of greenness dynamics within plausible estimates of photosynthetic carbon uptake. These model parameters were estimated by using long-term satellite observations of FAPAR (30 years), vegetation albedo (10 years), and mean annual GPP using a genetic optimization algorithm (chapter 3). LPJmL-GSI resulted in a better representation of seasonal, inter-annual (Figure 3.6) and decadal FAPAR dynamics (Figure 3.7). In particular, LPJmL-GSI reproduced observed greening and browning trends in the Sahel and in North American boreal forests respectively, which was not the case for LPJmL-OP. By considering observational

uncertainties from three satellite datasets and ten phenology detection methods, the results demonstrate that LPJmL-GSI outperforms LPJmL-OP and has an even better agreement with the satellite datasets than the datasets have with each other especially regarding the start and end of the growing season in boreal forests (Figures 4.3-4.5).

The implementation of the new phenology model and the optimization of model parameters additionally contributed to a better representation of global carbon cycle observations. LPJmL-GSI better reproduced tree cover and biomass especially in high-latitude ecosystems (Figure 3.8). LPJmL-GSI simulations of global carbon stocks and fluxes were closer to independent estimates than those of the original model (chapter 7.5.1). Furthermore, LPJmL-GSI simulations of NBP trends in high-latitude ecosystems agree well with independent estimates from atmospheric inversions (Figure 5.1, Figure 9.5). Finally, LPJmL-GSI reproduces the temporal dynamics of atmospheric CO<sub>2</sub> observations, especially of the amplitude of the seasonal cycle of CO<sub>2</sub>, which is an indicator for the photosynthetic carbon uptake of northern terrestrial ecosystems (Beer et al. (2010), chapter 5, Figures 5.1, 5.2, 9.7). In summary, DGVMs can be improved by a new generation of phenology models that simulate the continuous temporal development of canopies and that account for water availability. Such new phenology models in combination with parameter optimization approaches against long observational time series result in a better representation of seasonal to long-term greenness and carbon cycle dynamics.

The new phenology module in LPJmL can serve as a basis for future model developments and as a framework to assess alternative controls on vegetation phenology. Although FAPAR seasonality, growing season albedo, and mean annual GPP were optimized in LPJmL within this thesis, no attempt was made to directly improve the seasonal dynamics of ecosystem carbon and water fluxes. The results show that the seasonality of ecosystem carbon fluxes needs further improvement in LPJmL (in particular the seasonal peak photosynthetic carbon uptake happens too early in temperate PFTs, Figure 7.18). This model-data mismatch could be resolved by optimizing model parameters for photosynthesis, evapotranspiration and phenology against measurements of ecosystem carbon and water fluxes from eddy covariance observations (Carvalhais et al., 2010; Kuppel et al., 2012; Williams et al., 2009). Thereby the importance of additional controls on phenology that are currently not considered in

LPJmL-GSI could be explored. For example, chilling (i.e. a required time period of dormancy before leaves can develop) is of importance for some species (Migliavacca et al., 2012; Richardson et al., 2013) and leaf age might be an important factor for senescence globally (Caldararu et al., 2012, 2014). Although such controls might be of importance on individual or species-level, the results of this thesis show that such controls are not needed to explain global land surface phenology. Similar to the study of Migliavacca et al. (2012) the suitability of such alternative controls on phenology could be tested in a model-data integration framework with LPJmL-GSI but with the additional advantage of simultaneously modelling ecosystem carbon and water fluxes, stocks, and vegetation dynamics within a DGVM. Additionally to eddy covariance data, PhenoCam data and preferably additional measurements of leaf biomass and litter fall could be used to constrain parameters for productivity, phenology and leaf turnover. Such a model optimization against multiple ecosystem-level data streams would allow for the assessment of explanative capabilities of alternative controls on phenology and to accurately represent the relations between phenology, photosynthesis and evapotranspiration for ecosystems under a range of climate conditions.

Model-data integration of phenology and carbon and water fluxes at ecosystem level needs to be complemented by large scale observational constraints on the terrestrial carbon cycle to ensure the applicability of DGVMs for both regional and global-scale vegetation/carbon cycle simulations. Such constraints are, for example, long-term time series of atmospheric CO<sub>2</sub> measurements or newly available satellite datasets of carbon stocks like biomass. Satellite-derived biomass maps (Saatchi et al., 2011; Thurner et al., 2014) contain information about past vegetation dynamics like growth or mortality. By comparing modelled and observed biomass maps, model representations of vegetation turnover rates and the associated processes can be assessed (Carvalhais et al., 2014). Recent satellite-derived biomass maps and future retrievals from the “biomass” satellite mission (Le Toan et al., 2011) will potentially allow to estimate net changes in vegetation carbon stocks and to infer environmental mechanisms behind such changes. However, the suitability of satellite-derived biomass maps for model parameter optimization needs to be critically assessed with respect to the uncertainties of these datasets (Thurner et al., 2014). Although satellite-derived biomass maps probably

represent the spatial variability of carbon stocks well, it is not clear if they can provide constraints on model simulations of regional to global carbon cycle budgets. Therefore, long-term observations of atmospheric CO<sub>2</sub> provide a large-scale integrated constraint with less observational uncertainty. CO<sub>2</sub> time series are already used to estimate parameters of land surface models in global carbon-cycle data assimilation systems (CCDAS) (Kaminski et al., 2012, 2013). Nevertheless, CCDAS often have a coarse spatial resolution and thus provide only limited information about the impact of regional ecosystem and vegetation dynamics and their associated carbon balances on the global carbon cycle. On the one hand, atmospheric CO<sub>2</sub> observations can constrain large scale net carbon fluxes but only weakly constrain regional and gross carbon fluxes (Koffi et al., 2012). On the other hand, ecosystem-level model optimization experiments using eddy covariance flux data resulted in a good representation of the seasonality of ecosystem carbon fluxes (Kuppel et al., 2012; Stöckli et al., 2008a), but they do not necessarily improve the seasonality of atmospheric CO<sub>2</sub>, especially at mid-latitudes (Kuppel et al., 2014). Therefore it is recommended to use atmospheric CO<sub>2</sub> and ecosystem-level flux observations together to constrain CCDAS (Kaminski et al., 2013). This suggests applying model-data integration across spatial and temporal scales by using the respective scale-dependent data sets to constrain the corresponding parameters that are sensitive at this scale. Such a multi-scale/multi-data model optimization approach requires for example the development of methods that help to identify scale-dependent processes and model parameters (Mahecha et al., 2010b) and to optimally weigh a variety of data sets with different properties in parameter estimation (Wutzler and Carvalhais, 2014). The large and growing availability of observational data sets at different spatial scales requires the development of modelling and integration systems that are capable of representing ecosystem- to global-scale processes seamlessly, which will result in a better understanding of the Earth system.

### 6.3 Co-dominant water control on phenology and greenness dynamics

*The importance of environmental controls for land surface phenology and greenness varies regionally and depends on time scale. However, water availability and temperature play co-dominant roles globally for seasonal dynamics, inter-annual variability and trends in land surface phenology and greenness.*

On seasonal time scales, water availability is a major control of FAPAR dynamics not only in water-limited biomes but also in boreal forests and the arctic tundra (Figure 3.9). Although the seasonality of temperature and light are the dominant controls for the seasonality of greenness in temperate, boreal and arctic biomes, water availability was unexpectedly also of importance in permafrost-underlain boreal and arctic ecosystems. The seasonal freezing in autumn and thawing in spring of the upper layer of permafrost soils (the active layer) regulates water availability and thus contributes to leaf development in spring. Water availability was also of major importance for greenness seasonality in temperate grasslands where previous studies found stronger controls by temperature and light (Jolly et al., 2005). Moreover, water availability was the only control for seasonal greenness dynamics that was of importance in all biomes, whereas the importance of temperature and light was limited to regions north of 40°N or high altitudes (Figure 3.9).

Inter-annual variability and trends of land surface phenology in high-latitude ecosystems are controlled by light and water availability in addition to temperature (chapter 4.4.2). The warming-induced trend of an advancing start of season in temperate and boreal ecosystems is limited by the availability of light. Therefore future warming trends in high-latitude ecosystems might not necessarily result in a prolonged growing season. Additionally, water availability and permafrost had a strong effect on trends in the start of growing season trends in Siberian boreal forests (Figure 4.8). This result suggests that melting permafrost results in higher water availability (at least in lowlands (Shur and Jorgenson, 2007)) which supports earlier spring green-up of vegetation.

Inter-annual variability and trends of peak greenness were controlled by regionally different combinations of factors but water availability and land use and land cover change were important factors across all regions (Figure 4.6a, chapter 1.4.3). The greening of the tundra and of boreal forests was mostly affected by temperature effects



on phenology, from land cover change and permafrost dynamics. Increasing temperatures result in a prolongation of the growing season, melting permafrost and the resultant higher water availability, which together contribute to increased productivity and vegetation coverage. Boreal forests in North America are less underlain by permafrost than boreal forests in Eastern Siberia. Consequently, the role of permafrost was less important in boreal forests in North America than in Siberia which contributes to the divergent pattern of less greening/more browning in boreal North America compared to Siberia. In temperate forests in Europe, land use/land cover change was the dominant control on greening besides the effect of temperature on phenology. Greening in Europe is often related to a re-growth of forests after socio-economic changes in central and eastern Europe (Alcantara et al., 2013; Kuemmerle et al., 2008; Main-Knorn et al., 2009). In the Sahel, seasonal effects of water availability and land use/land cover change were the dominant controls on greening. CO<sub>2</sub> fertilization contributed globally to greening but was of very minor importance in comparison to climate effects on phenology or land use/land cover change. Fire disturbance had an important effect on inter-annual variability and trends in vegetation greenness regionally (savannahs, parts of boreal forests). In summary, seasonal dynamics, inter-annual variability and trends of land surface phenology and greenness can be regionally explained by different sets of controlling factors but water availability is a co-dominant control in all regions – and therefore globally as well.

Several other controls on greenness dynamics, or more generally on plant growth and vegetation dynamics, exist. For example, plant productivity and growth is limited by the availability of the macronutrients N, P, S, K, Ca, and Mg, and of the trace elements Fe, Mn, Zn, Cu, Mo, B, and Cl (Larcher, 2003, p.196). Such dependencies of plant growth on nutrients are currently not implemented in LPJmL. Plant productivity is limited by nitrogen availability across the globe (LeBauer and Treseder, 2008) and especially in high-latitude ecosystems (Vitousek and Howarth, 1991). Whereas a decreased availability of nitrogen results in a browning and loss of needles (Larcher, 2003, p.201), increased nitrogen availability results in an increased productivity of forests ecosystems (Hyvönen et al., 2007). Consequently, the increasing availability of nitrogen through deposition or through the release from thawing permafrost soils in northern ecosystem can increase plant productivity (Keuper et al., 2012) and thus

might be an important control on greening. However, another modelling study suggests that changes in nitrogen deposition were of minor importance for greenness trends in comparison to climatic changes in northern ecosystems (Mao et al., 2012). Nevertheless, the assessment of the role of nitrogen limitation on global productivity trends under changing environmental conditions is challenging: On the one hand because of a poor representation of nitrogen uptake processes in ecosystem models, and on the other hand because of the limited availability and understanding of observations (Zaehle et al., 2014). Consequently, responses of plant productivity and growth to nutrient availability (nitrogen and beyond) need to be implemented or improved in all ecosystem models and compared to a growing number of well documented and suitable observations or experiments.

Several other controls affect greenness trends on regional scales. For example, industrial air pollution caused vegetation degradation around Norilsk in northern Siberia (Tutubalina and Rees, 2001), and in the so called “Black Triangle” regions between Germany, Poland, the Czech Republic and Slovakia (Kubíková, 1991; Main-Knorn et al., 2009; Markert et al., 1996; Schulze, 1989). However, the reduction of heavy industry after 1990 and the introduction of air pollution mitigation policies decreased emissions in eastern-central Europe. This resulted in a regional re-growth of forests, but it highly depends on forest management and disturbance histories (Main-Knorn et al., 2009). Nevertheless, the Black Triangle region turned into a “Green Triangle” (Ladysz, 2006). These regions experienced the strongest greening trends in Europe (Figure 1.2). However it is not clear if the increase in NDVI is really due to increased vegetation activity or due to the reduced aerosol content in the atmosphere for which the GIMMS NDVI3g dataset is only marginally corrected (Pinzon and Tucker, 2014). Nevertheless, these regional findings demonstrate the influence of human activities on ecosystems and regional biogeography through biogeochemical processes. Other disturbances such as insect outbreaks strongly affect forest productivity and structure (Wermelinger, 2004) and might become more important under future climate warming (Hlásny et al., 2011; Jönsson et al., 2009). Indeed regional observations demonstrate a co-occurrence of insect infestations with satellite-observed browning trends (Buma et al., 2013; Neigh et al., 2014; Olsson et al., 2012). Insects outbreaks can strongly affect regional carbon balances by turning forests from carbon sinks to sources which might

be of similar or even greater importance than the effect of wild fires (Kurz et al., 2008). However, the long-term development of net ecosystem productivity is highly sensitive to outbreak conditions and to the post-outbreak forest management and succession (Edburg et al., 2011, 2012). In summary, regional natural or human-induced disturbances need to be considered to explain greenness dynamics but their peculiarities make the consideration of such processes in global ecosystem models a challenging task for future research.

#### **6.4 Vegetation dynamics as regulator of carbon cycle trends**

*The increasing seasonal amplitude of atmospheric CO<sub>2</sub> in northern latitudes is mainly driven by an enhanced terrestrial gross primary production in boreal and arctic ecosystems, which is caused by the climate-induced greening trend.*

Greening in boreal and arctic ecosystems is related to an increase in deciduous vegetation and to positive trends in gross primary production (Figures 6.1, 9.5). Trends in photosynthetic carbon uptake were stronger than trends in ecosystem respiration which results in increasing net biome productivity. By coupling LPJmL with the TM3 transport model it was possible to quantify the effect of this increased net uptake of high-latitude ecosystems on the seasonal amplitude of atmospheric CO<sub>2</sub>. LPJmL was able to reproduce the latitudinal gradient of stronger CO<sub>2</sub> amplitude trends in northern latitudes than in low latitudes (Figures 6.2a, 9.4). The fertilizing effect of rising atmospheric CO<sub>2</sub> on photosynthesis and agricultural land use change contributed to CO<sub>2</sub> amplitude trends, but these factors cannot explain the much stronger trends in northern latitudes. Climate change and greening in boreal and arctic regions were the dominant controls on the trend in the seasonal CO<sub>2</sub> amplitude (Figure 6.2). According to these results the “agricultural green revolution” (Zeng et al., 2014) and the “direct human influence (...) from increased cropland production” (Gray et al., 2014) on the trend in the seasonal CO<sub>2</sub> amplitude were overestimated because these studies did not comprehensively account for other controlling factors such as increased vegetation activity in boreal and arctic ecosystems and the role of atmospheric transport.

Uncertainties of future trends in atmospheric CO<sub>2</sub> are related to differences between models in the simulation of terrestrial carbon cycle and vegetation dynamics

(Friedlingstein et al., 2014). Processes of vegetation dynamics like turnover of plant compartments, mortality, disturbances, intra- and inter-specific competition affect vegetation composition and finally control the residence time of carbon in terrestrial vegetation (Friend et al., 2014). These processes dominate the uncertainty in future carbon stocks and fluxes in the terrestrial vegetation. However, improving dynamic global vegetation models to better represent such processes of vegetation dynamics like tree mortality is challenging because of a lack of empirical data and process understanding (Allen et al., 2010; Anderegg et al., 2013; McDowell et al., 2011). For example, a globally consistent documentation of forest mortality does not exist; the often species-specific physiological processes of mortality are poorly understood; there are no appropriate global maps of vegetation change at a level of species or functional groups; and mortality interacts with climate and disturbances like fire and especially insects for which also no appropriate global observations exist (Allen et al., 2010; Hartmann et al., 2015). Consequently, it is currently “impossible to conclude what mechanisms and level of detail are needed” (McDowell et al., 2011, p.8) to improve processes of mortality in vegetation models. However, the curse of mortality events is often species-specific, which is challenging to represent in most of the current dynamic global vegetation models that rely on a limited set of plant functional types. Therefore the implementation of individual-dependent plant traits in ecosystem models (Bodegom et al., 2014; Pavlick et al., 2013; Sakschewski et al., 2015) could support the representation of the peculiarities of mortality events. Thereby trait-based modelling approaches can be constraint by relations between traits as observed in global plant trait databases (Kattge et al., 2011; Madani et al., 2014; Niinemets et al., 2015). Alternative approaches could make use of international tree ring databases (Barichivich, 2014; Rammig et al., 2015) or of satellite-derived biomass maps (Saatchi et al., 2011; Thurner et al., 2014) to infer past productivity, environmental stress, mortality and vegetation dynamics. Biomass maps together with global estimates of ecosystem carbon fluxes (Jung et al., 2011) allow the estimation of ecosystem or vegetation turnover times (Carvalhais et al., 2014). Refining and compartmentalization of such estimates of turnover could overcome the current lack of understanding and could contribute to a better understanding of mortality in ecosystem carbon turnover.

In summary, the uncertainty of future climate/carbon cycle interactions results from a limited present-day understanding. The understanding of the role of vegetation dynamics in the Earth system can be expanded by an enhanced availability of global observations and observation-based estimates and model-data integration approaches. Observations at multiple spatial-temporal scales – e.g. tree rings and near-surface remote sensing at the tree level; eddy covariance measurements at the ecosystem level; satellite-derived greenness, land cover change, and biomass at the regional level; in-situ or satellite-based CO<sub>2</sub> measurements at the continental to global level – need to be integrated within vegetation-atmosphere models to evaluate and improve our current codified theories of Earth system functioning, and to consistently quantify Earth system dynamics. Model-data integration is therefore the key for a consistent understanding of the Earth system.

---

What is science? Humboldt or Gauß?

## 7 Supplement of chapter 3

---

This chapter contains the Supplement of the publication (chapter 3):

Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

Supplement of *Biogeosciences*, 11, 7025–7050, 2014  
<http://www.biogeosciences.net/11/7025/2014/>  
doi:10.5194/bg-11-7025-2014-supplement  
© Author(s) 2014. CC Attribution 3.0 License.



*Supplement of*

**Identifying environmental controls on vegetation greenness phenology through model–data integration**

M. Forkel et al.

## 7 Supplement of chapter 3

### 7.1 LPJmL model details

#### 7.1.1 Original phenology model (LPJmL-OP)

The phenology model in the original LPJmL formulation has three different routines for summergreen (i.e. temperature-driven deciduous), evergreen (no seasonal variation) and rain-green (i.e. water-driven deciduous) PFTs (Sitch et al., 2003). Evergreen PFTs have a constant phenology status ( $Phen = 1$ ). The daily phenology status of summergreen PFTs depends on growing degree-days (GDD):

$$\begin{aligned}\Delta T &= T - GDD_{base} \\ GDD_t &= GDD_{t-1} + \Delta T_t \quad \text{if } \Delta T > 0\end{aligned}\tag{7.1}$$

Where  $T$  is the daily air temperature and  $GDD_{base}$  is the minimum temperature threshold to start counting GDDs. Daily GDD is scaled to the phenology status using a parameter  $ramp$  which is the amount of GDDs to get full leaf cover:

$$Phen_{PFT|summergreen} = \begin{cases} GDD / ramp & \text{if } aphen < aphen_{max} \\ 0 & \text{if } aphen \geq aphen_{max} \\ 0 & \text{if } aphen > aphen_{min} \text{ and } \Delta T < 0 \end{cases}\tag{7.2}$$

The daily phenology status is set back to 0 if the accumulated phenology status ( $aphen$ ) is larger than a parameter  $aphen_{max}$  or if  $aphen$  is greater than  $aphen_{min}$  and the daily temperature is below  $GDD_{base}$ . The daily accumulated phenology status is calculated as:

$$aphen_t = aphen_{t-1} + Phen_t\tag{7.3}$$

For rain-green PFTs the daily phenology status is calculated dependent on the daily water availability scaling factor  $Wscal$  in LPJmL (Supplement 7.2) (Gerten et al., 2004) and a threshold value ( $Wscal_{min}$ ):

$$Phen_{PFT|raingreen} = \begin{cases} 1 & \text{if } Wscal \geq Wscal_{min} \\ 0 & \text{if } Wscal < Wscal_{min} \end{cases}\tag{7.4}$$

The phenology of rain-green PFTs has no smooth behaviour but is a binary switch between full leaf cover and no leaves according to this formulation. For herbaceous PFTs the same phenology scheme like for summergreen PFTs is used but the phenology status is only set back to 0 at the end of the phenology year (i.e. on the 14<sup>th</sup>

day of the year for the northern hemisphere and on the 195<sup>th</sup> day of the year for the southern hemisphere).

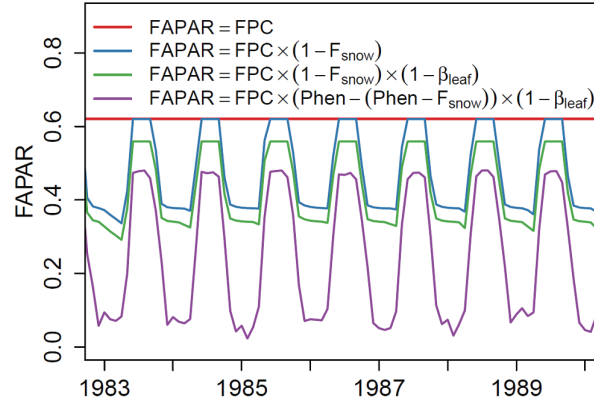


Figure 7.1: Effects on FAPAR in LPJmL for an example grid cell in Siberia.

FAPAR in LPJmL is computed from foliar projective cover (FPC), from snow coverage in the green canopy ( $F_{snow}$ ), leaf albedo ( $\beta_{leaf}$ ) and phenology status (Phen).

### 7.1.2 Water availability scaling factor

The water availability scaling factor  $W_{scal}$  in LPJmL is a ratio between water supply  $S$  and atmospheric water demand  $D$  for a dry canopy (Gerten et al., 2004):

$$W_{scal} = \frac{S}{D} \quad (7.5)$$

In the LPJmL-GSI phenology model the water availability scaling factor is expressed as a percentage value:

$$W = W_{scal} \times 100 \quad (7.6)$$

Water supply is dependent on the maximum transpiration  $E_{max}$  under water saturation and relative soil moisture  $w_r$  (Gerten et al., 2004):

$$S = E_{max} \times w_r \quad (7.7)$$

Atmospheric water demand  $D$  for a dry canopy is calculated from potential evapotranspiration  $PET$ , maximum Priestley-Taylor coefficient  $\alpha_{max} = 1.391$ , scaling canopy conductance  $g_m = 3.26 \text{ mm s}^{-1}$  and potential canopy conductance  $g_{pot}$  (Gerten et al., 2004):

$$D = \frac{PET \times \alpha_{max}}{1 + (g_m / g_{pot})} \quad (7.8)$$



### 7.1.3 Albedo

Surface albedo and snow coverage routines have been implemented in LPJmL to use it as a land surface scheme in a coupled vegetation-climate model (Strengers et al., 2010). We used this implementation but made the albedo parameters PFT-dependent as albedo differs between ecosystems (Cescatti et al., 2012). The albedo of a grid cell  $Alb_{gc}$  is the area-weighted sum of the vegetation albedo  $Alb_{veg}$ , bare-soil albedo  $Alb_{bare}$  and snow albedo:

$$Alb_{gc} = Alb_{veg} + F_{bare} \times (F_{snow} \times \beta_{snow} + (1 - F_{snow}) \times \beta_{soil}) \quad (7.9)$$

where  $F_{bare}$  and  $F_{snow}$  are the coverage of bare soil and snow on top of bare soil in a grid cell and  $\beta_{soil}$  and  $\beta_{snow}$  are the soil and snow albedo parameters, respectively. The parameters  $\beta_{soil} = 0.4$  and  $\beta_{snow} = 0.7$  were used as constants (Strengers et al., 2010) and not further considered in this study. Although soil and snow albedo has clear spatial and temporal variations which are due to changing moisture contents, an improvement of these processes is not within the scope of our study. The vegetation albedo is computed as the albedo of each PFT  $Alb_{PFT}$  and its corresponding FPC:

$$Alb_{veg} = \sum_{PFT=1}^{PFT=n} Alb_{PFT} \times FPC_{PFT} \quad (7.10)$$

The albedo of a PFT depends on the fraction of the PFT that is completely covered by snow  $F_{snow,PFT}$  and the albedo of the PFT without snow coverage ( $Alb_{PFT,nosnow}$ ) (Strengers et al., 2010):

$$Alb_{PFT} = F_{snow,PFT} \times \beta_{snow} + (1 - F_{snow,PFT}) \times Alb_{PFT,nosnow} \quad (7.11)$$

The albedo of a PFT without snow coverage is the sum of leaf, stem/branches and litter (background) albedo:

$$Alb_{PFT,nosnow} = Alb_{leaf,PFT} + Alb_{stem,PFT} + Alb_{litter,PFT} \quad (7.12)$$

The albedo of green leaves depends on the foliar projective cover, the daily phenology status and the PFT-dependent leaf albedo parameter:

$$Alb_{leaf,PFT} = FPC_{PFT} \times Phen_{PFT} \times \beta_{leaf,PFT} \quad (7.13)$$

The albedo of stems and branches depends on the fractional coverage of the ground by stems and branches ( $cstem$ ) and a PFT-dependent stem albedo parameter  $\beta_{stem,PFT}$ :

$$Alb_{stem,PFT} = FPC_{PFT} \times (1 - Phen_{PFT}) \times cstem \times \beta_{stem,PFT} \quad (7.14)$$

The parameter  $c_{stem} = 0.7$  (Strengers et al., 2010) was used as a constant and not further considered in this study. The background (i.e. litter) albedo of a PFT depends additionally on a PFT-dependent litter albedo parameter  $\beta_{litter,PFT}$ :

$$Alb_{litter,PFT} = FPC_{PFT} \times (1 - Phen_{PFT}) \times (1 - c_{stem}) \times \beta_{litter,PFT} \quad (7.15)$$

The parameters  $\beta_{leaf,PFT}$ ,  $\beta_{stem,PFT}$  and  $\beta_{litter,PFT}$  were implemented as PFT-dependent albedo parameters which differs from the previous implementation (Strengers et al., 2010). The fraction of snow in the green part of the canopy that is used to compute FAPAR (equation 3) depends on the daily phenological status and the fraction of the PFT that is covered by snow:

$$F_{snow,gv,PFT} = Phen_{PFT} \times F_{snow,PFT} \quad (7.16)$$

The fraction of the PFT that is covered by snow depends on snow height and the daily calculated snow water equivalent (Strengers et al., 2010).

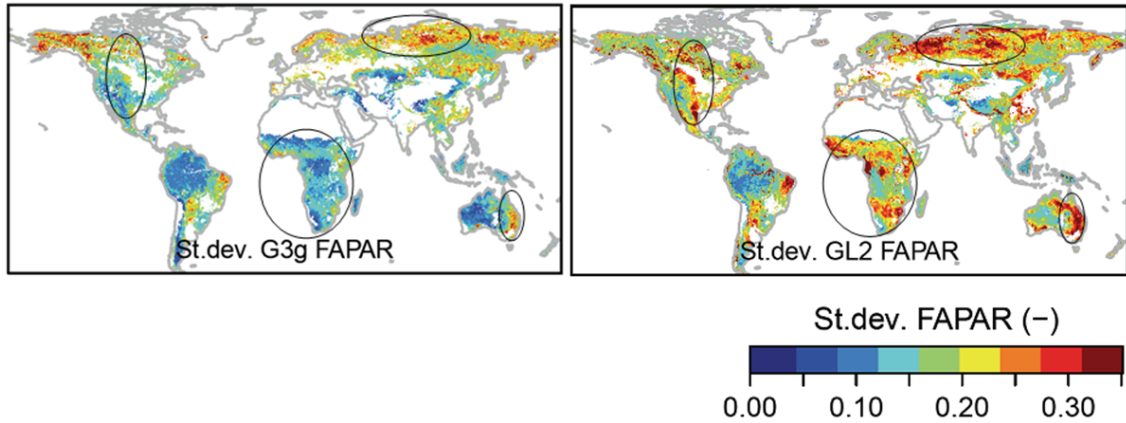


Figure 7.2: Standard deviation of mean annual FAPAR from the GIMMS3g and GL2 FAPAR datasets in 1982-2011.

The annual mean FAPAR was calculated for each year from each monthly FAPAR value for months with monthly mean air temperatures  $> 0^{\circ}\text{C}$ . Areas with large differences are highlighted with circles.

## 7.2 FAPAR datasets

### 7.2.1 Comparison of the Geoland2 and GIMMS3g FAPAR datasets

We compared the Geoland2 and GIMMS3g FAPAR datasets to assess 1) the agreement of two newly developed FAPAR products and 2) to evaluate the suitability of these products for the optimization of FAPAR and phenology-related parameters in LPJmL. We found important differences between the Geoland2 and GIMMS3g FAPAR datasets

during our analyses. The differences are mostly related to inter-annual variability and trends.

The GL2 FAPAR dataset had a higher inter-annual variability in most regions especially in northern Russia, central North America, Africa and eastern Australia (Figure 7.2). Despite the different amplitudes of inter-annual variability, the temporal dynamic of annual aggregated FAPAR values was well correlated in most regions (Figure 7.25). Nevertheless, in some regions like in the North American Tundra, in parts of the Siberian boreal forest and in the tropical forests the inter-annual temporal FAPAR dynamic was weakly or even negatively correlated (Figure 7.1).

The temporal dynamics of mean annual FAPAR agreed relatively well between GIMMS3g FAPAR and GL2 FAPAR in the AVHRR period. The temporal dynamic of mean annual FAPAR agreed poorly between GIMMS3g and GL2 FAPAR in the VGT period. Both datasets had higher biases in boreal needle-leaved evergreen forests (Figure 7.3). An offset between the GL2 AVHRR and GL2 VGT FAPAR time series in the overlapping years 1999 and 2000 is evident in all biomes. Additionally, the GL2 VGT time series shows an abrupt jump from 2002 to 2003 which is probably due to the sensor change from VGT1 to VGT2 (Horion et al., 2014). Because of these reasons, the Geoland2 FAPAR dataset cannot be used for a long-term analysis of FAPAR trends and extremes.

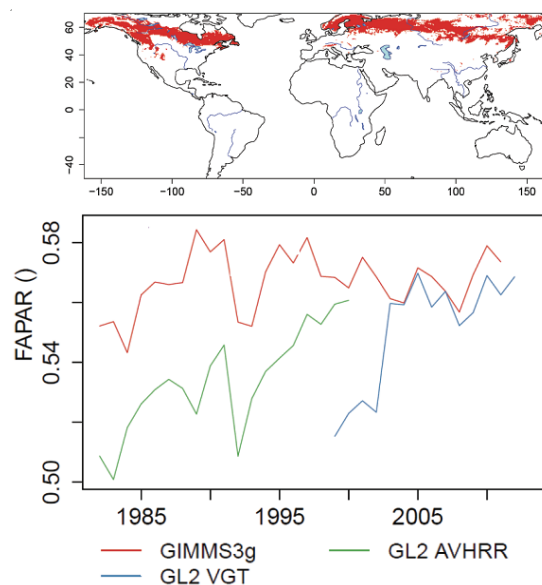


Figure 7.3: Comparison of mean annual FAPAR from different datasets averaged for the extent of boreal needle-leaved evergreen forests.

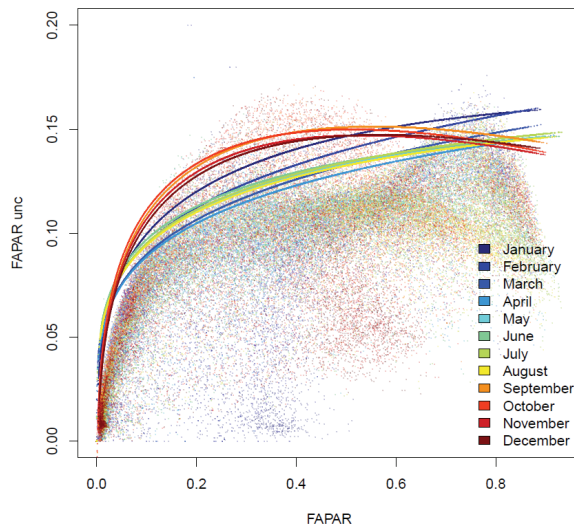


Figure 7.4: Monthly quantile regressions between GL2 VGT FAPAR and the GL2 VGT FAPAR fitted to the quantile 0.95.

Each monthly quantile regression was applied to the GIMMS3g FAPAR dataset to estimate uncertainties for this dataset. Using 0.95 quantile regressions provides conservative uncertainty estimates for the GIMMS3g FAPAR dataset.

### 7.2.2 Estimation of uncertainty for the GIMMS3g FAPAR dataset

The GIMMS3g FAPAR dataset was used for parameter optimization. For parameter optimization it is necessary to consider data uncertainty in multiple data stream cost functions. Unfortunately, the GIMMS3g dataset has no uncertainty estimates. On the other hand the GL2 FAPAR dataset has uncertainty estimates but time series are not well harmonized. Thus we were using the GIMMS3g dataset for parameter optimization but estimated uncertainties by using regression to the uncertainty of the GL2 FAPAR dataset (Figure 7.4). Therefore we fitted for each month polynomial quantile regressions to the quantile 0.95 between FAPAR and FAPAR uncertainty from the GL2 VGT FAPAR dataset. Then we were using these regressions to estimate uncertainties for the GIMMS3g FAPAR dataset.

## 7.3 Land cover

### 7.3.1 Creation of an observation-based map of plant functional types

Land cover maps from remote sensing products are not directly comparable with plant functional types in global vegetation models because they are using different legends for the description of vegetation (Jung et al., 2006; Poulter et al., 2011a). Land cover

classes have to be reclassified into the corresponding PFTs. We were using the SYNMAP land cover map (Jung et al., 2006), the Köppen-Geiger climate classification (Kottek et al., 2006) and tree coverage from MODIS (Townshend et al., 2011). We decided to use the SYNMAP land cover map because it offers fractional land coverage and synergizes already the GLCC, MODIS and GLC2000 land cover maps (Jung et al., 2006). PFTs in LPJmL are defined according to biome (tropical, temperate or boreal), leaf type (needle leaved, broadleaved) and phenology (summergreen, evergreen, rain green). We extracted the biome information from the Köppen-Geiger climate classification whereas leaf type and phenology were extracted from the SYNMAP land cover map. The FPC of a PFT was derived from MODIS tree cover.

In a first step, we reclassified the Köppen-Geiger climate classification in to bioclimatic zones (biomes) that correspond to the definition used in LPJmL (Figure 7.5). This reclassification followed to a large extent the rules of Poulter et al. (2011a):

- The climate zone A was reclassified to the tropical biome.
- The climate regions BWh and BSh were reclassified to the tropical biome.
- The climate regions BWk and BSk were reclassified to the temperate biome.
- The climate region Cw was reclassified to the tropical biome.
- The climate regions Cf and Cs were reclassified to the temperate biome.
- The climate regions D and E were reclassified to the boreal biome.

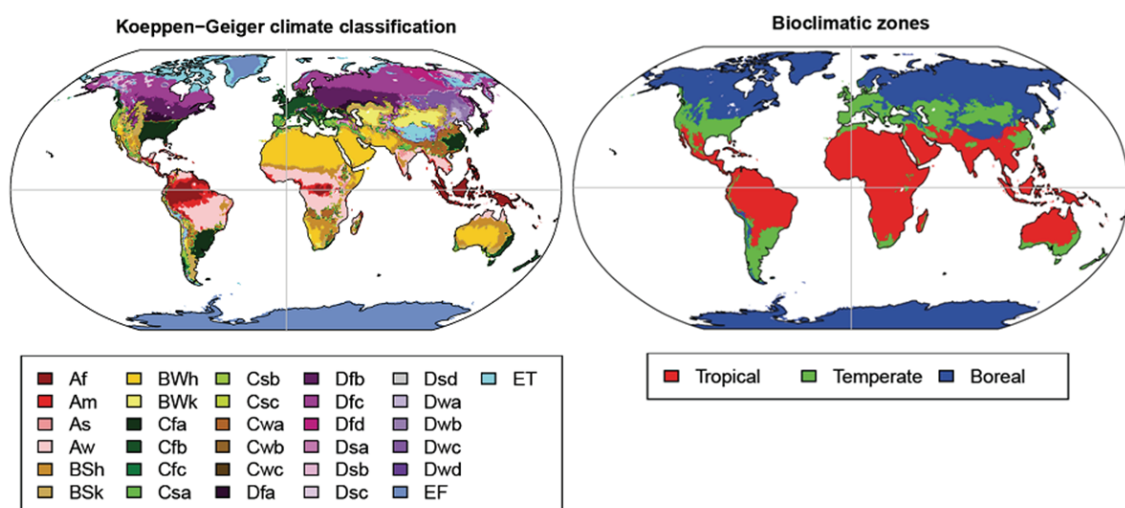


Figure 7.5: Reclassification of the Koeppen-Geiger climate classification in bioclimatic zones.

In a second step, we created a land cover map with PFT legend by crossing the land cover information from SYNMAP with the map of biomes following rules for each tree PFT:

- TrBE: EBF (evergreen broadleaved forest) AND tropical biome
- TrBR: DBF (deciduous broadleaved forest) AND tropical biome
- TeNE: ENF (evergreen needleleaved forest) AND temperate biome
- TeBE: EBF (evergreen broadleaved forest) AND temperate biome
- TeBS: DBF (deciduous broadleaved forest) AND temperate biome
- BoNE: ENF (evergreen needleleaved forest) AND boreal biome
- BoBS: DBF (deciduous broadleaved forest) AND boreal biome
- BoNS: DNF (deciduous needleleaved forest) AND boreal biome

Although we translated in this step the land cover classes into PFTs, the fractions represent still fraction of land cover and not FPC. For example, a grid cell can be covered by 100% forest but this forest contains only 70% trees while the rest is covered by herbaceous plants. This difference becomes evident by comparing the total coverage of forest land cover classes from SYNMAP with tree cover from MODIS (Figure 7.6). MODIS tree cover is always lower than forest cover but shows more spatial variability. In a third step, we need to correct the land cover fraction with tree cover to create a map of FPC. Thus, we calculated the FPC of each tree PFT by correcting the land cover fraction of a PFT ( $LC_{PFT}$ ) with the ratio of fractional tree coverage from MODIS ( $F_{Tree}$ ) and the total land coverage of all 8 forest PFTs:

$$FPC_{PFT} = LC_{PFT} \times \frac{F_{Tree}}{\sum_{PFT=1}^{PFT=8} LC_{PFT}} \quad (7.17)$$

This calculation of FPC differs from the approach of Poulter et al. (2011a) who divided each land cover class in fixed fractions of tree and herbaceous PFTs.

In the last step we need to calculate the FPC of herbaceous PFTs:

$$FPC_{herb} = 1 - F_{Tree} - LC_{Barren} - LC_{Water} - LC_{Snow/Ice} \quad (7.18)$$

which is the residual area by removing the fractional tree coverage from MODIS and the land cover fractions of bare soil and rocks, water and permanent snow and ice from the total grid cell. Thus, grasslands, croplands and shrub lands were assigned to

herbaceous vegetation. Then we divided the herbaceous FPC into the TeH, PoH and TrH PFTs according to biomes:

- TrH:  $FPC_{herb}$  AND tropical biome
- Old TeH:  $FPC_{herb}$  AND temperate OR boreal biome

The TeH was further splitted in a new temperate herbaceous and a polar herbaceous PFT to separate between temperate grasslands and tundra:

- TeH (new): old TeH AND temperate OR boreal biome AND boreal trees  $< 0.3$
- PoH: old TeH AND (boreal biome OR Koeppen-Geiger E climate) AND boreal trees  $> 0.3$

These steps yielded in observation-based maps of foliar projective cover for each PFT (Figure 7.7). As the input data (SYNMAP and MODIS VCF) is based on satellite data from the years 2000/2001 the retrieved maps reflect the distribution of PFTs for the year 2000.

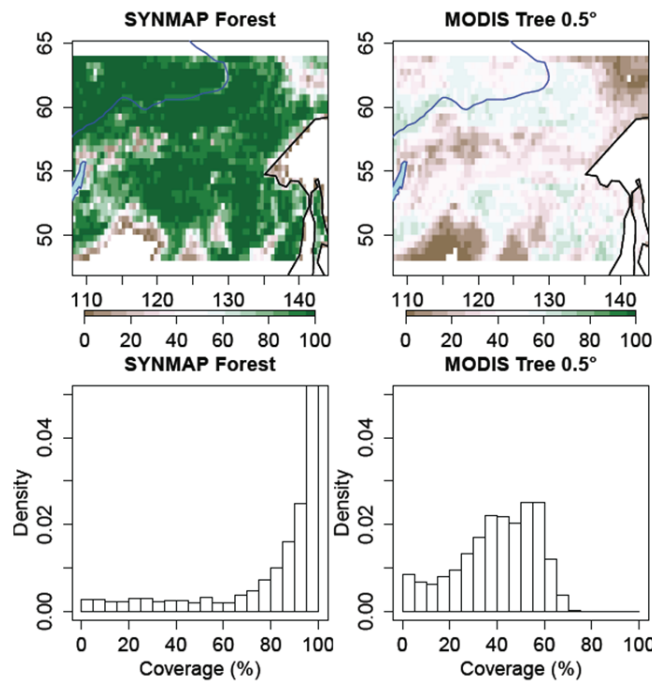


Figure 7.6: Comparison of total forest coverage from SYNMAP and MODIS tree coverage for a region in eastern Siberia.

### 7.3.2 Comparison of simulated and observed PFT distributions

We compared the observation-based PFT map with the simulated PFT distribution from LPJmL-OP for the year 2000. LPJmL with dynamic vegetation simulated usually too high tree and too low herbaceous cover in all regions (Figure 7.8). In the central

tropical forests (Amazon, Congo basin) LPJmL simulated too low cover of TrBE but too high cover of TrBR. The coverage of BoNE was too low in some regions in North America and Eastern Siberia. The simulated distribution of BoNS did not agree much with the observed distribution which is almost limited to eastern Siberia. Tree cover was especially overestimated in regions with only sparse tree cover (Savannahs, Steppe/boreal forest transition, eastern Siberia). The extent of boreal forest PFTs (BoNE, BoBS, BoNS) is generally too large with far southward extensions into the Steppe and northward extensions into the Tundra.

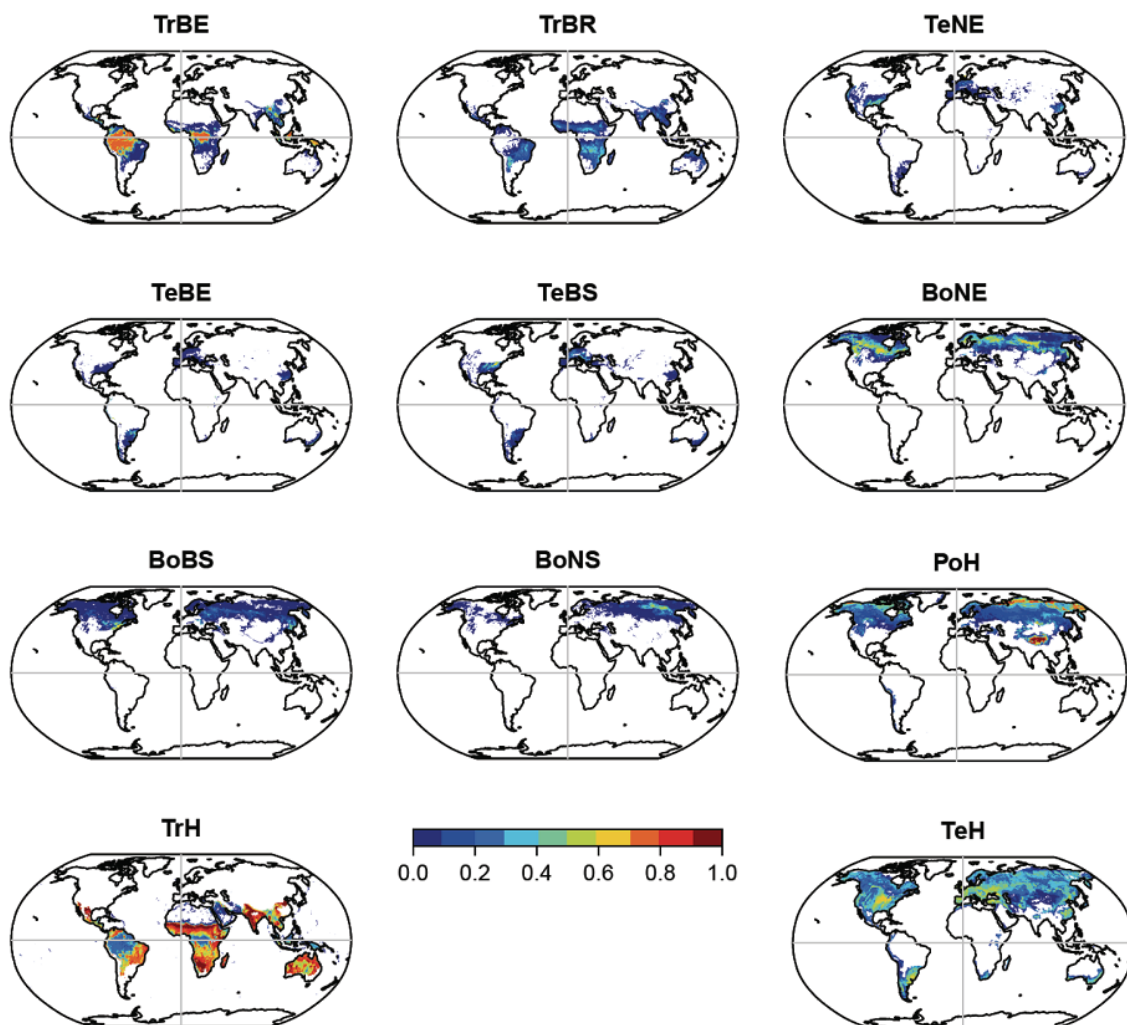


Figure 7.7: Observation-based maps of the foliar projective cover of plant functional types. Agricultural areas are included in the TrH and TeH PFTs.

As expected, the prescription of the observed PFT maps into LPJmL generally improved the representation of the observed PFT distributions (Figure 7.8). The spatial patterns of PFT distributions were highly correlated and the bias in comparison to the



observed distribution was clearly reduced in comparison with the model run with dynamic vegetation. The PFT distribution of the LPJmL model run with prescribed land cover does not perfectly agree with the observed PFT distribution which is due to the applied prescription approach. Tree PFTs can have a lower FPC in LPJmL than the prescribed FPC value because the trees are still growing or because mortality reduced the FPC. This effect especially happened in the BoNE PFT where fire reduced the FPC in large regions in Canada and eastern Siberia (Figure 7.8). Herbaceous PFTs can have a higher FPC than the observed FPC value because these PFTs were allowed to establish the entire grid cell (except the fraction that is barren, water or permanent snow/ice in the observations). This happened for example when fires burnt tree PFTs and herbaceous PFTs succeeded afterwards in LPJmL. This is the reason for the overestimation of herbaceous coverage in large regions in Canada and eastern Siberia where the BoNE PFT was underestimated (Figure 7.8). In summary, the prescription of land cover improved the representation of observed spatial patterns of PFTs in LPJmL. Differences to the observed PFT distribution are due to the desired ability of LPJmL to represent important processes of vegetation dynamics like mortality processes.

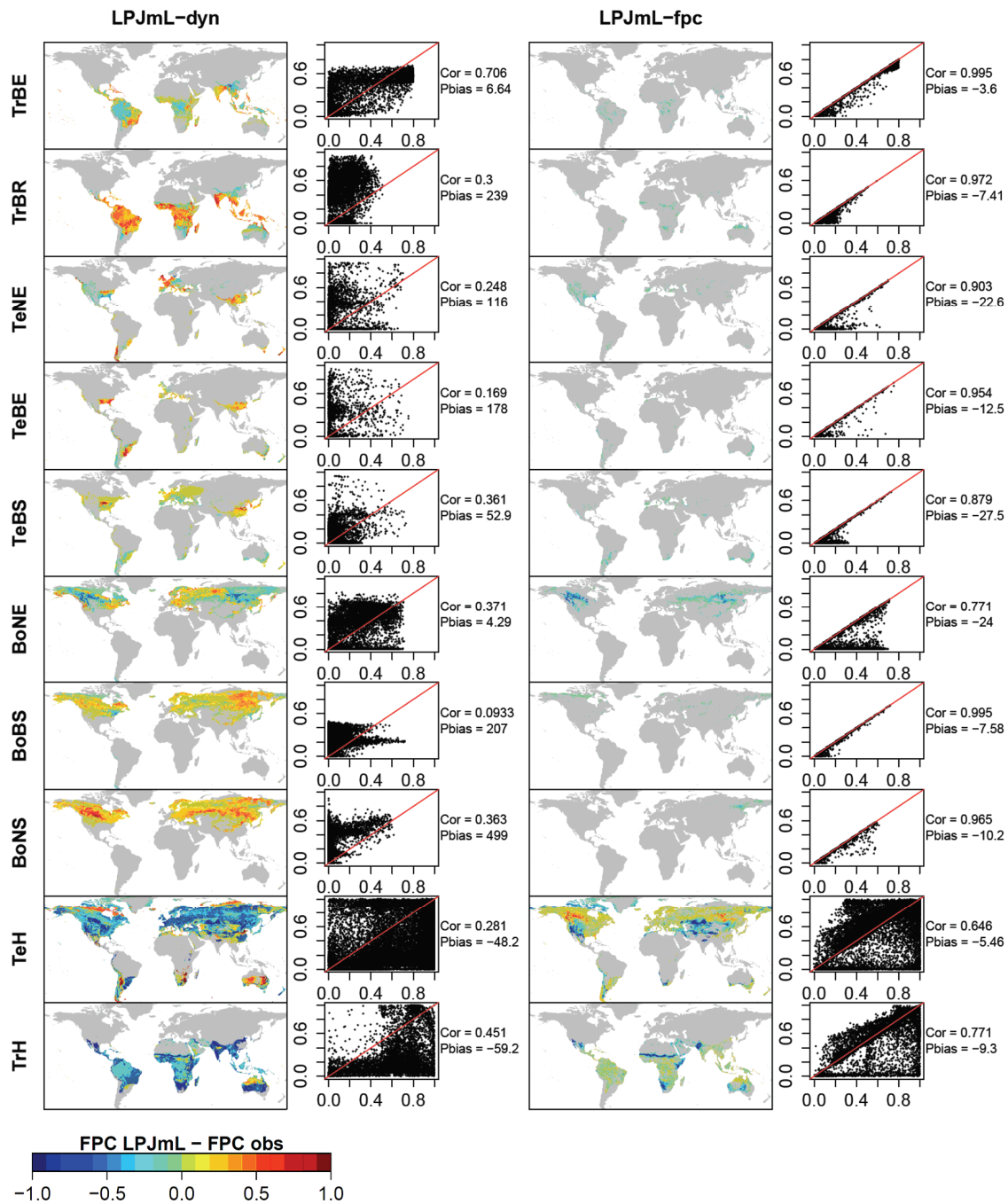


Figure 7.8: Comparison between simulated and observed PFT distributions for the year 2000. The maps are difference maps between simulated FPC values from LPJmL-OP and observed FPC values. The scatter plots show observed FPC values on the x-axis and simulated FPC values on the y-axis. Left: LPJmL-OP with dynamic vegetation and prescribed burnt areas. Right: LPJmL-OP with prescribed land cover and prescribed burnt areas.

## 7.4 Model parameter optimization

### 7.4.1 Parameter definitions and values

This section documents the LPJmL parameters that were addressed in this study. The parameters and their use in the model are described in Table 7.1. The information

sources from which prior parameter values were extracted for each optimization experiment are shown in Figure 7.9. Tables 7.2-5 list prior and posterior parameter values of each optimization experiment according to the logical flow of optimization experiments indicated in Figure 7.9.

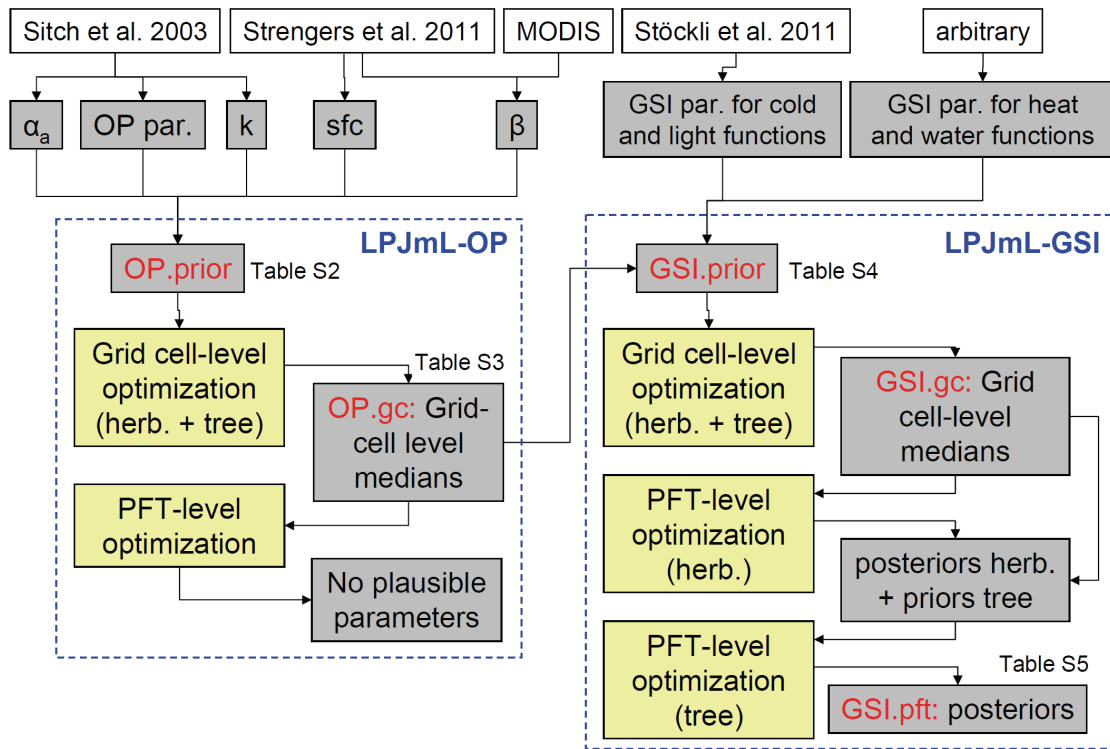


Figure 7.9: Information sources for prior and posterior parameter sets and overview of model optimization experiments.

Grey boxes indicate model parameters or parameter sets. White boxes are information sources for parameters. Yellow boxes are optimization experiments.

Table 7.1: Description of LPJmL model parameters that were addressed in this study.

Parameter	Alternative name	Use	Description	Unit
$\alpha_a$	ALPHAA	Photo-synthesis	Leaf-to-canopy scaling parameter (amount of radiation absorbed at leaf-level in comparison to total canopy)	-
$\beta_{leaf}$	ALBEDO_LEAF	Albedo, FAPAR	Albedo of green leaves	-
$\beta_{stem}$	ALBEDO_STEM	Albedo	Albedo of stems and branches	-
$\beta_{litter}$	ALBEDO_LITTER	Albedo	Albedo of litter	-
k	LIGHTTEXTCOEFF	FPC, FAPAR	Light extinction coefficient in Lambert-Beer relationship	-
sfc	SNOWCANOPYFRA C	Albedo, FAPAR	Maximum fraction of snow in the green canopy	-
$W_{scal_{min}}$	MINWSCAL	Original phenology	Minimum value of the water availability scaling factor for leaf onset in rain green PFTs	-
$GDD_{base}$	GDDBASE	Original phenology	Minimum daily temperature to start counting growing degree days	$^{\circ}C$
ramp	RAMP	Original phenology	Number of growing degree days to reach full leave cover in summergreen PFTs	$^{\circ}C$
$aphen_{min}$	APHEN_MIN	Original phenology	Minimum accumulated phenology state to allow senescence if temperature < GDDBASE	-
$aphen_{max}$	APHEN_MAX	Original phenology	Maximum accumulated phenology state. Phenology is set back to 0 if this value is passed.	-
$sl_{tmin}$	TMIN_SL	GSI phenology	Slope of cold temperature limiting logistic function for phenology	$1/^{\circ}C$
$base_{tmin}$	TMIN_BASE	GSI phenology	Inflection point of cold temperature limiting logistic function for phenology	$^{\circ}C$
$T_{tmin}$	TMIN_TAU	GSI phenology	Change rate of actual to previous day cold temperature limiting function value for phenology	-
$sl_{light}$	LIGHT_SL	GSI phenology	Slope of light limiting logistic function for phenology	$1/(W/m^2)$
$base_{light}$	LIGHT_BASE	GSI phenology	Inflection point of light limiting logistic function for phenology	$W/m^2$
$T_{light}$	LIGHT_TAU	GSI phenology	Change rate of actual to previous day light limiting function value for phenology	-
$sl_{water}$	WATER_SL	GSI phenology	Slope of water limiting logistic function for phenology	1/%
$base_{water}$	WATER_BASE	GSI phenology	Inflection point of water limiting logistic function for phenology	%
$T_{water}$	WATER_TAU	GSI phenology	Change rate of actual to previous day water limiting function value for phenology	-
$sl_{heat}$	TMAX_SL	GSI phenology	Slope of heat limiting logistic function for phenology	$1/^{\circ}C$
$base_{heat}$	TMAX_BASE	GSI phenology	Inflection point of heat limiting logistic function for phenology	$^{\circ}C$
$T_{heat}$	TMAX_TAU	GSI phenology	Change rate of actual to previous day heat limiting function value for phenology	-

Table 7.2: Prior parameter values of LPJmL-OP (OP.prior).

The values in brackets are ranges of uniform parameter distributions that were used during optimization. Note: \* The parameter GDDbase was changed to 0°C. This value gave better agreements between simulated and observed seasonal FAPAR dynamics than the original value of 5°C. Nevertheless, GDDbase was not included in optimization experiments because this parameter is highly correlated with the parameter ramp.

	TrBE	TrBR	TeNE	TeBE	TeBS	BoNE	BoBS	BoNS	TeH	TrH
$\alpha_a$	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)
$\beta_{leaf}$	0.15 (0.1-0.2)	0.15 (0.1-0.2)	0.15 (0.06-0.23)	0.15 (0.09-0.23)	0.16 (0.086-0.23)	0.14 (0.05-0.23)	0.15 (0.09-0.21)	0.12 (0.1-0.15)	0.14 (0.072-0.22)	0.15 (0.09-0.21)
$\beta_{stem}$	0.15 (0.018-0.29)	0.15 (0.073-0.23)	0.13 (0-0.31)	0.15 (0.029-0.28)	0.13 (0.038-0.23)	0.14 (0-0.31)	0.14 (0.059-0.23)	0.13 (0.052-0.32)	--	--
$\beta_{litter}$	0.15 (0.018-0.29)	0.14 (0.058-0.27)	0.13 (0.047-0.21)	0.15 (0.044-0.29)	0.14 (0.085-0.2)	0.13 (0.035-0.26)	0.14 (0.078-0.22)	0.12 (0.088-0.23)	0.14 (0.027-0.38)	0.13 (0.02-0.28)
sfc	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)	0.4 (0.1-0.9)
k	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)	0.5 (0.1-0.9)
GDD <sub>ba</sub>	--	--	--	--	0*	--	0*	0*	0*	0*
W <sub>scal</sub>	--	0.3 (0-1)	--	--	--	--	--	--	--	--
Ramp	--	--	--	--	300 (0-1000)	--	200 (0-1000)	200 (0-1000)	100 (0-1000)	100 (0-1000)
aphen <sub>min</sub>	--	--	--	--	10 (1-600)	--	10 (1-600)	10 (1-600)	--	--
aphen <sub>max</sub>	--	--	--	--	210 (1-600)	--	210 (1-600)	210 (1-600)	--	--

#### 7.4.2 Genetic optimization algorithm

We were using a genetic optimization algorithm to minimize the cost function  $J(d)$  by optimizing the scaled parameter vector  $d$ . The GENOUD algorithm (genetic optimization using derivatives) (Mebane and Sekhon, 2011) combines global genetic optimization search with local gradient-based search algorithms. In genetic optimization algorithms, each model parameter is called a gene and each parameter set is called an individual. The fitness of this individual is the cost of the model against the observations. At the beginning of the optimization, a first generation of individuals is initialized by random sampling of parameter sets within the prescribed parameter

ranges. After the calculation of the cost of all individuals of the first generation, a next generation is generated by cloning the best individuals, by mutating the genes or by crossing different individuals (Mebane and Sekhon, 2011). This results after some generations in a set of individuals with highest fitness, i.e. parameter sets with minimized cost. Within the GENOUD algorithm we were using also the BFGS (Broyden-Fletcher-Goldfarb-Shanno) gradient search algorithm (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970) to find an optimum parameter set. An optimized parameter set of the BFGS algorithm is used as individual in the next generation. The BFGS gradient search algorithm was first applied on the best individual of the second last generation to avoid a too fast convergence of the optimization algorithm towards a local optimum. For grid cell-based optimization experiments we were applying the GENOUD algorithm with at least 20 generations and a population size of 1000 individuals per generation, i.e. at least 20000 single model runs. For PFT-level optimization experiments we were applying the GENOUD algorithm with at least 15 generations and a population size of at least 700 individuals per generation, i.e. at least 10500 single model runs.

Table 7.3: Posterior parameter values for LPJmL-OP based on grid cell-level optimization experiments (OP.gc).

Parameters written *in italics* were derived as the median value of the single grid cell optimization experiments whereas all other parameters were derived from prior parameter sources. For the parameter ramp no plausible parameter was found. The parameter GDDbase was changed to 0 but not included in the optimization.

	<b>TrBE</b>	<b>TrBR</b>	<b>TeNE</b>	<b>TeBE</b>	<b>TeBS</b>	<b>BoNE</b>	<b>BoBS</b>	<b>BoNS</b>	<b>TeH</b>	<b>TrH</b>
$\alpha_a$	<i>0.6</i>	<i>0.56</i>	<i>0.38</i>	<i>0.41</i>	<i>0.38</i>	<i>0.28</i>	<i>0.34</i>	<i>0.27</i>	<i>0.32</i>	<i>0.39</i>
$\beta_{\text{leaf}}$	<i>0.13</i>	<i>0.1</i>	<i>0.06</i>	<i>0.1</i>	<i>0.16</i>	<i>0.05</i>	<i>0.18</i>	<i>0.11</i>	<i>0.08</i>	<i>0.15</i>
$\beta_{\text{stem}}$	0.15	<i>0.07</i>	0.13	0.15	<i>0.04</i>	0.14	<i>0.06</i>	<i>0.05</i>	--	--
$\beta_{\text{litter}}$	0.15	<i>0.06</i>	0.13	0.15	<i>0.09</i>	0.13	<i>0.08</i>	<i>0.09</i>	<i>0.1</i>	<i>0.14</i>
sfc	0.4	0.4	<i>0.1</i>	0.4	0.4	0.1	<i>0.15</i>	<i>0.18</i>	0.4	0.4
k	<i>0.36</i>	<i>0.73</i>	<i>0.41</i>	<i>0.44</i>	<i>0.74</i>	<i>0.71</i>	<i>0.51</i>	<i>0.88</i>	<i>0.39</i>	<i>0.46</i>
GDD <sub>ba</sub> <sub>se</sub>	--	--	--	--	0	--	0	0	0	0
Wscal <sub>min</sub>	--	<i>0.85</i>	--	--	--	--	--	--	--	--
Ramp <sub>min</sub>	--	--	--	--	300	--	200	200	100	100
aphen <sub>min</sub>	--	--	--	--	10	--	10	10	--	--
aphen <sub>max</sub>	--	--	--	--	<i>201.9</i> 7	--	<i>181.6</i> 2	<i>105.7</i> 8	--	--

Table 7.4: Prior parameter values for LPJmL-GSI (GSI.prior).

Parameters marked with \* were identified as insensitive and were not included in the optimization. The values in brackets are ranges of uniform parameter distributions that were used during optimization. The values for the first 6 parameters were derived from the single grid-cell optimization experiments of LPJmL-OP (Table 7.3).

	TrBE	TrBR	TeNE	TeBE	TeBS	BoNE	BoBS	BoNS	TrH	TeH PoH
$\alpha_a$	0.6 (0.2-0.8)	0.56 (0.1-0.9)	0.38 (0.23-0.49)	0.41 (0.1-0.9)	0.38 (0.15-0.6)	0.28 (0.16-0.57)	0.34 (0.15-0.61)	0.27 (0.16-0.55)	0.39 (0.21-0.83)	0.32 (0.1-0.83)
$\beta_{leaf}$	0.13 (0.1-0.2)	0.1 (0.05-0.2)	0.06 (0.01-0.23)	0.1 (0.09-0.23)	0.16 (0.13-0.19)	0.05 (0.01-0.23)	0.18 (0.09-0.21)	0.11 (0.1-0.14)	0.15 (0.09-0.21)	0.08 (0.072-0.22)
$\beta_{stem}$	0.15 (0.018-0.29)	0.07 (0.06-0.23)	0.13 (0-0.31)	0.15 (0.029-0.28)	0.04 (0.038-0.23)	0.14 (0-0.31)	0.06 (0.059-0.23)	0.05 (0.04-0.32)	--	--
$\beta_{litter}$	0.15 (0.054-0.29)	0.06 (0.058-0.27)	0.13 (0.047-0.21)	0.15 (0.044-0.29)	0.09 (0.085-0.2)	0.13 (0.035-0.26)	0.08 (0.078-0.22)	0.09 (0.088-0.23)	0.14 (0.02-0.28)	0.1 (0.027-0.38)
sfc	0.4*	0.4*	0.1 (0.01-0.9)	0.4*	0.4 (0.1-0.9)	0.1 (0.01-0.9)	0.15 (0.1-0.9)	0.18*	0.4*	0.4 (0.1-0.9)
k	0.36 (0.2-0.9)	0.73 (0.1-0.9)	0.41 (0.1-0.9)	0.44 (0.1-0.9)	0.74 (0.1-0.9)	0.71 (0.1-0.9)	0.51 (0.1-0.9)	0.88 (0.1-0.9)	0.46 (0.1-0.9)	0.39 (0.1-0.9)
$sl_{tmin}$	0.24 (0.1-2)	0.24*	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)	0.24 (0.1-2)
$base_{tmin}$	8.8 (0-16)	8.8 (0-16)	-3.3 (-6-6)	-0.6 (-3-1)	7.4 (5-9)	3.7 (-6-6)	2.2 (0-5)	-4 (-6-6)	8.8 (0-16)	0.7 (-3-5)
$T_{tmin}$	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*
$sl_{heat}$	0.24 (0.01-3)	0.24 (0.01-3)	0.24 (0.01-3)	0.24 (0.01-3)	0.24 (0.01-3)	0.24*	0.24*	0.24*	0.24 (0.01-3)	0.24*
$base_{heat}$	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)	35 (25-45)
$T_{heat}$	0.2 (0.01-0.9)	0.2*	0.2 (0.01-0.9)	0.2*	0.2*	0.2*	0.2 (0.01-0.9)	0.2*	0.2 (0.01-0.9)	0.2*
$sl_{light}$	57 (0.05-157)	23*	20*	0.2 (0.05-40)	58*	14*	101 (0.05-220)	95*	41 (0.05-130)	23*
$base_{light}$	125 (1-200)	62 (1-200)	73 (1-200)	23 (1-50)	123 (50-200)	57 (1-100)	166 (50-200)	156 (130-180)	104 (1-150)	67 (1-180)
$T_{light}$	0.2 (0.01-0.9)	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2*	0.2 (0.01-0.9)	0.2 (0.01-0.9)
$sl_{water}$	5 (0.1-10)	5 (0.1-10)	5*	5*	5 (0.1-10)	5*	5 (0.1-10)	5*	5 (0.1-10)	5 (0.1-10)
$base_{water}$	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)	20 (1-99)
$T_{water}$	0.8 (0.01-0.99)	0.8 (0.01-0.99)	0.8*	0.8*	0.8*	0.8*	0.8*	0.8*	0.8 (0.01-0.99)	0.8 (0.01-0.99)

Table 7.5: Final parameters for LPJmL-GSI.

Parameters written *in italics* were derived from PFT-level optimization experiments (GSI.pft) whereas all other parameters were derived from prior parameter sources as described in Figure 7.9.

	TrBE	TrBR	TeNE	TeBE	TeBS	Bo NE	Bo BS	Bo NS	TrH	TeH	PoH
$\alpha_a$	<i>0.63</i>	<i>0.52</i>	<i>0.44</i>	<i>0.45</i>	<i>0.61</i>	0.22	0.41	0.34	0.40	0.32	0.43
$\beta_{\text{leaf}}$	0.13	<i>0.12</i>	<i>0.12</i>	<i>0.12</i>	<i>0.18</i>	0.10	0.16	0.12	0.24	0.18	0.07
$\beta_{\text{stem}}$	0.10	0.10	0.04	0.04	0.04	0.06	0.06	0.04	0.15	0.15	0.15
$\beta_{\text{litter}}$	0.10	0.10	<i>0.05</i>	0.10	<i>0.14</i>	<i>0.01</i>	<i>0.00</i>	<i>0.01</i>	0.12	<i>0.07</i>	0.03
k	<i>0.52</i>	<i>0.74</i>	<i>0.47</i>	<i>0.70</i>	<i>0.60</i>	<i>0.44</i>	0.41	<i>0.66</i>	0.50	0.50	0.50
$sl_{\text{tmin}}$	1.01	0.24	0.22	0.55	0.26	0.10	0.22	0.15	0.91	0.31	0.13
$base_{\text{tmin}}$	8.30	7.66	<i>-7.81</i>	<i>-0.63</i>	<i>13.69</i>	<i>-7.52</i>	2.05	<i>-4.17</i>	6.42	<i>4.98</i>	2.79
$T_{\text{tmin}}$	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	<i>0.01</i>	0.20
$sl_{\text{heat}}$	1.86	<i>1.63</i>	1.83	0.98	1.74	0.24	1.74	0.24	1.47	0.24	0.24
$base_{\text{heat}}$	<i>38.64</i>	<i>38.64</i>	<i>35.26</i>	<i>41.12</i>	<i>41.51</i>	<i>27.32</i>	<i>41.51</i>	<i>44.60</i>	<i>29.16</i>	<i>32.04</i>	<i>26.12</i>
$T_{\text{heat}}$	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
$sl_{\text{light}}$	77.17	23.00	20.00	18.83	58.00	14.00	58.00	95.00	64.23	23.00	23.00
$base_{\text{light}}$	<i>55.53</i>	<i>13.01</i>	<i>4.87</i>	<i>39.32</i>	<i>59.78</i>	<i>3.04</i>	<i>59.78</i>	<i>130.1</i>	69.90	<i>75.94</i>	<i>50.00</i>
$T_{\text{light}}$	0.52	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.40	0.22	0.38
$sl_{\text{water}}$	5.14	7.97	5.00	5.00	5.24	5.00	5.24	5.00	0.10	0.52	0.88
$base_{\text{water}}$	<i>5.00</i>	<i>22.21</i>	<i>8.61</i>	<i>8.82</i>	<i>20.96</i>	<i>0.01</i>	<i>20.96</i>	<i>2.34</i>	<i>41.72</i>	<i>53.07</i>	<i>1.00</i>
$T_{\text{water}}$	0.44	<i>0.13</i>	0.80	0.80	0.80	0.80	0.80	0.80	0.17	0.01	0.94

### 7.4.3 Parameter sensitivities and uncertainties

To explore the sensitivity and uncertainty of LPJmL-GSI parameters after PFT-level optimizations, we computed the likelihood  $L$  and Akaikes Information Criterion AIC from the cost  $J$  of each individual (i.e. parameter set  $d$ ) of the genetic optimization:

$$L = e^{-J(d)} \quad (7.19)$$

$$AIC = 2 \times n - 2 \times \log(L) \quad (7.20)$$

Where  $n$  is the number of parameters. The optimum parameter set has the highest likelihood and the lowest AIC. Then, we selected only these individuals with an AIC difference  $dAIC$  of  $< 2$  in comparison to the best parameter set:

$$dAIC = AIC - AIC_{\text{best}} \quad (7.21)$$

Parameter sets or model formulations with an AIC difference  $< 2$  are usually considered as equally plausible like the best parameter set (Burnham and Anderson, 2002, p.70). The relationship between likelihood and the value of each parameter provides both a qualitative insight in the uncertainty of parameters as expressed by the



parameter range and in the parameter sensitivity as expressed by the maximum likelihood at each parameter value.

#### 7.4.4 Supplementary results and discussion on optimization performance

The optimization of LPJmL-OP and LPJmL-GSI resulted in a significant reduction of the cost in comparison to the respective prior models although there were differences between plant functional types (Figure 7.10). LPJmL-OP with prior parameters had high costs especially in herbaceous PFTs (TrH and TeH) and in the boreal needle-leaved summer green PFT (BoNS). The optimization of single grid cells in LPJmL-OP resulted in a significant reduction of the cost in all PFTs ( $p \leq 0.01$ , Wilcoxon rank-sum test) despite the polar herbaceous and tropical herbaceous PFTs. The global prior parameter set of LPJmL-GSI resulted in a significant lower cost than the grid cell-level optimized parameter sets of LPJmL-OP in TrH, TeBS, BoNS and PoH PFTs. The optimization of single grid cells in LPJmL-GSI resulted in a significant reduction of the cost in all PFTs except BoNS and PoH. PFT-level optimizations of LPJmL-GSI resulted in a significant lower cost than the LPJmL-GSI prior parameter set in all PFTs except TeBE, BoNS and PoH. PFT-level optimizations of LPJmL-GSI resulted in a significant lower cost than the standard LPJmL-OP prior parameter set in all PFTs except TeNE. These results demonstrate an improved overall performance of optimized model parameter sets over prior model parameter sets and of LPJmL-GSI over LPJmL-OP regarding a cost that is defined based on 30 years of monthly FAPAR, mean annual GPP and 10 years of monthly vegetation albedo.

Model optimization experiments resulted in a significant reduction of the annual GPP bias of LPJmL in comparison to the MTE data-oriented GPP product (Figure 7.11). LPJmL-OP with prior parameters underestimated mean annual GPP in the TrBE PFT (median Pbias -13%) and overestimated mean annual GPP in all other PFTs (up to 123% median Pbias in TeH). Grid cell-level optimization experiments of LPJmL-OP resulted in a significant reduction of the GPP bias in all PFTs except in the PoH PFT. Especially in the TrBE, TrBR, TrH, TeNE, TeBE, TeBS and BoBS PFTs the bias of mean annual GPP of LPJmL was removed almost completely (i.e. Pbias within 5%). The LPJmL-GSI prior parameter set had significant lower biases of mean annual GPP than the prior parameter set of LPJmL-OP. This was because the median of each parameter

from the OP.gc experiments was used as prior parameter for LPJmL-GSI. Grid cell-level optimization experiments of LPJmL-GSI resulted in significant reductions of the bias in mean annual GPP in most PFTs despite PFTs where the LPJmL-GSI prior parameter set resulted already in GPP biases close to 0 (i.e. TrH, TeBE and PoH). PFT-level optimization experiments of LPJmL-GSI resulted in significant lower biases of mean annual GPP than the prior parameter set of LPJmL-OP in all PFTs except PoH. These results demonstrate that through the applied model optimization biases in mean annual GPP were significantly reduced in all PFTs (except PoH) in LPJmL-OP as well as in LPJmL-GSI.

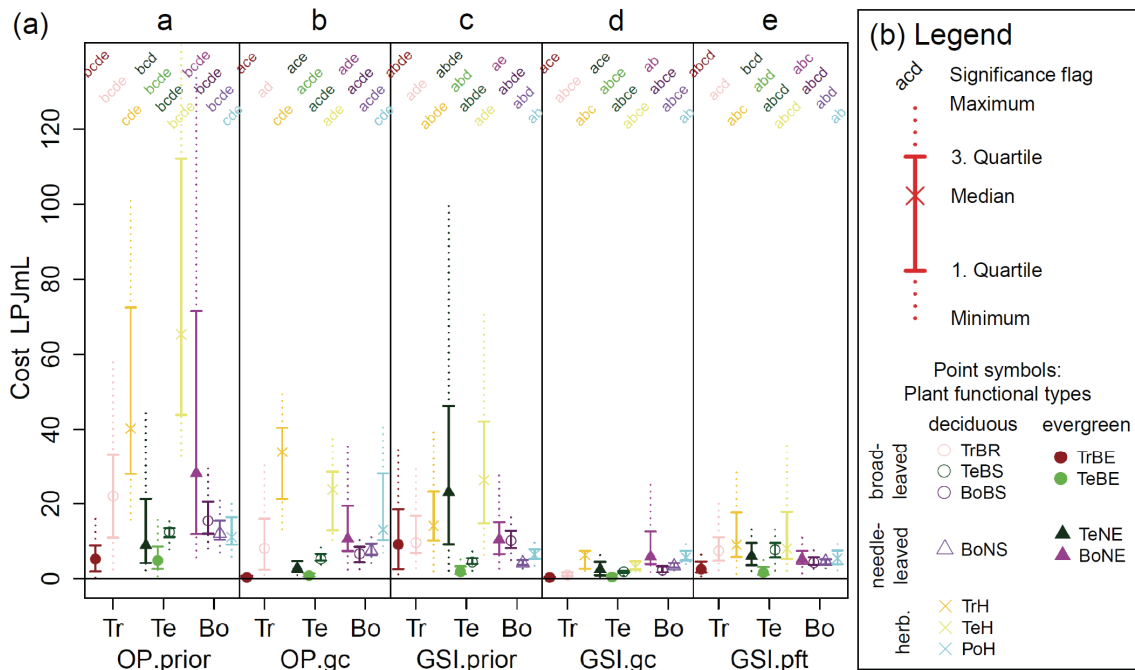


Figure 7.10: Distribution of the cost for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes.

(a) Cost for LPJmL-OP with default parameters (a, OP.prior), after grid cell-level optimizations (b, OP.gc), cost for LPJmL-GSI with prior parameters (c, GSI.prior), after grid cell-level optimizations (d, GSI.gc) and after PFT-level optimizations (e, GSI.pft). Biomes are Tr (tropical), Te (temperate) and Bo (boreal/polar).

(b) Legend for the plot. Each distribution is plotted according to usual boxplot statistics. The point symbols indicate the plant functional type. The significance flag on top of each distribution shows if a distribution is significant different ( $p \leq 0.01$ ) to the corresponding distribution of the same PFT in another optimization experiment.

The significance is based on the Wilcoxon rank-sum test. For example “acd” indicates a significant difference to the main categories a (OP.prior), c (GSI.prior) and d (GSI.gc) but no significant difference to b (OP.gc) and e (GSI.pft).

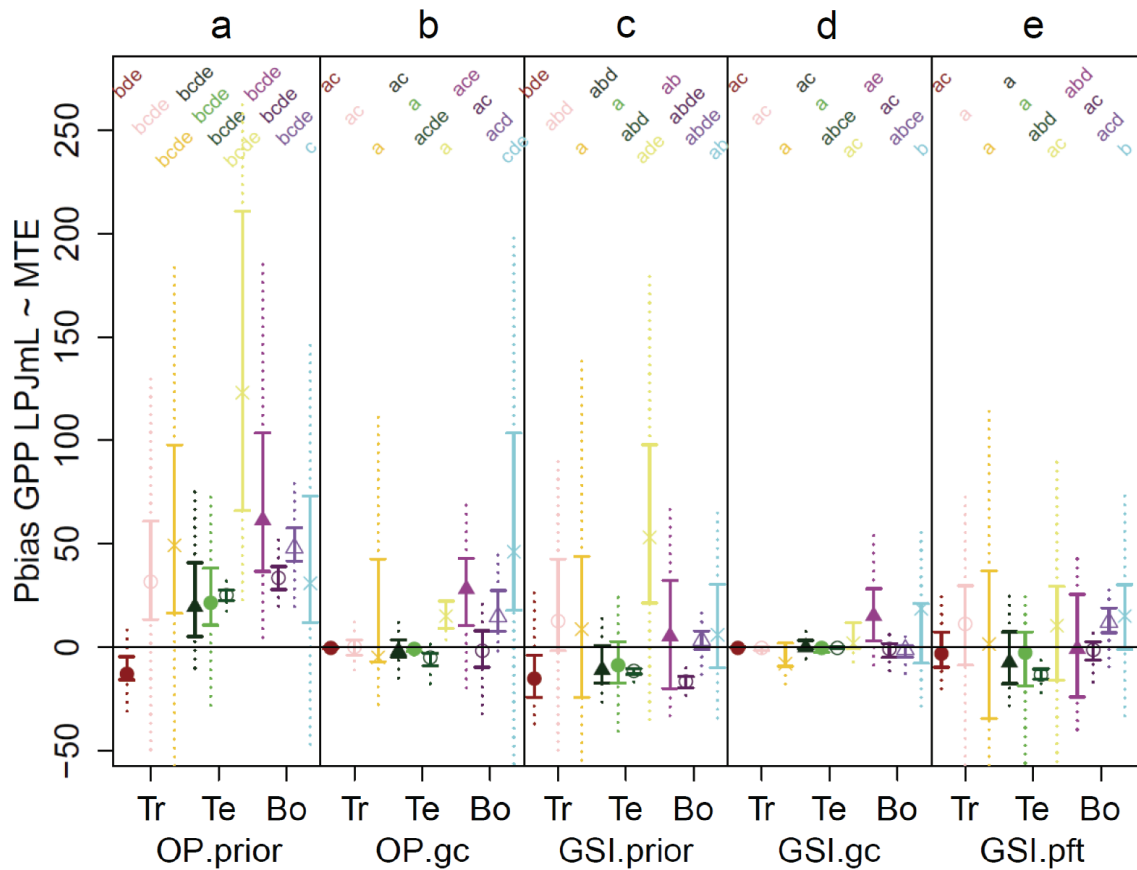


Figure 7.11: Distribution of the percent bias between LPJmL and MTE mean annual GPP (1982-2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes. See Figure 7.10 for a further explanation of this figure.

We were not able to remove the GPP bias and to reduce the cost of LPJmL-OP and of LPJmL-GSI in the PoH PFT (tundra) in optimization experiments because of inconsistencies between the FAPAR and GPP datasets or in the LPJmL formulation. Although a complete removal of the GPP bias is in principle possible by adjusting the  $\alpha_a$  parameter, this would result in a too low FPC of the PoH PFT. Such a low FPC cannot explain the relatively high peak FAPAR values that are seen in the GIMMS3g FAPAR dataset in Tundra regions. It is not possible to explain the low mean annual MTE GPP and the relatively high GIMMS3g peak FAPAR with the current LPJmL model structure in tundra regions. The reasons for this mismatch can be caused by inconsistencies between the GPP and FAPAR datasets or by an insufficient model formulation. The MTE data-oriented GPP product has been upscaled from FLUXNET eddy covariance measurements (Jung et al., 2011). Nevertheless, not many eddy covariance measurement sites cover tundra regions with mean annual air temperatures  $< 0^\circ\text{C}$ . Thus, the MTE GPP estimates are not well supported by measurements in

tundra regions. But also the FAPAR dataset might be more uncertain in tundra regions than in other parts of the globe. Optical remote sensing in high-latitude regions is usually performed under high-sun zenith angles. Radiation can penetrate deeper into vegetation under high-sun zenith angles which results in higher FAPAR (Tao et al., 2009; Walter-Shea et al., 1998). Thus, the high FAPAR values in the GIMMS3g FAPAR dataset might be caused by satellite observations under high-sun zenith angles. Finally, the inconsistencies between GPP and FAPAR might be also caused by an inappropriate representation of tundra plant communities in LPJmL. The PoH PFT in LPJmL was derived from a grass PFT but does not include shrubs or the large functional diversity of mosses and lichen that are the dominant plant communities in tundra ecosystems (Porada et al., 2013). We currently cannot decide if the inconsistency between FAPAR and GPP in our optimization of productivity and FAPAR parameters in tundra regions is more caused by the specific properties of the datasets or by an insufficient model structure.

All optimization experiments resulted in reasonable albedo biases of LPJmL-OP and LPJmL-GSI in comparison with monthly MODIS albedo time series (Figure 7.12). LPJmL-OP with prior parameters overestimated growing season albedo in all PFTs. Grid cell-level optimization experiments of LPJmL-OP resulted in significant reductions of the bias in growing season albedo in TrBE, TeNE, TeBE, TeBS, BoNE, and BoNS PFTs but not in TrBR, TrH, TeH, BoBS and PoH PFTs. The bias in growing season albedo of the latter PFTs was significantly reduced with the LPJmL-GSI prior parameter set. The optimization of LPJmL-GSI for single grid cells significantly reduced the bias in growing season albedo in comparison to the LPJmL-GSI prior parameter set in all PFTs except in the TeH, BoNS and PoH PFTs. These results demonstrate that model optimizations experiments kept growing season albedo within reasonable ranges in comparison to MODIS albedo.

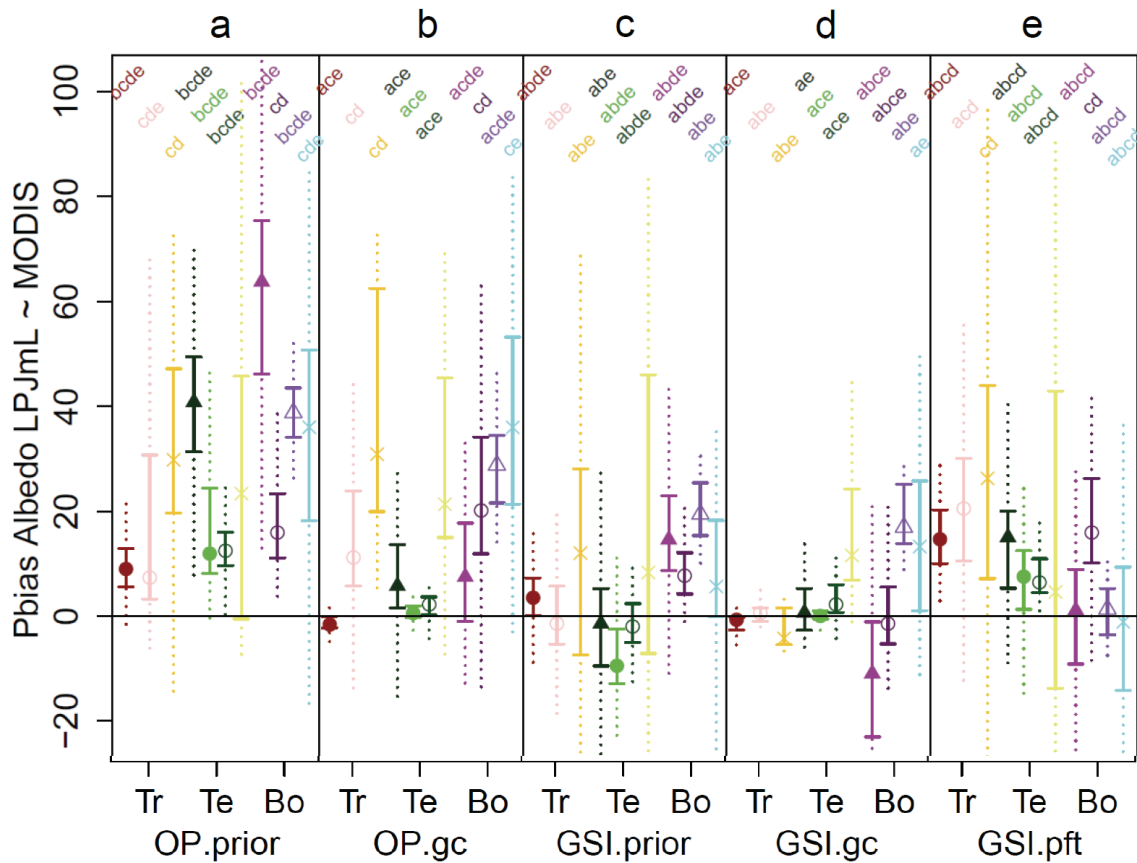


Figure 7.12: Distribution of the percent bias between LPJmL and MODIS monthly growing season albedo (2000-2011) for several grid cells in prior model runs and optimization experiments grouped by plant functional types and biomes. See Figure 7.10 for a further explanation of this figure.

#### 7.4.5 Supplementary results and discussion on parameter variability

The optimization of the leaf albedo parameter  $\beta_{\text{leaf}}$  resulted in values that differed especially between broadleaved and needle-leaved evergreen PFTs (Figure 7.13). Needle-leaved evergreen PFTs (TeNE and BoNE) had in all optimization experiments the lowest  $\beta_{\text{leaf}}$  parameter values while the broad-leaved summergreen PFTs (TeBS and BoBS) had the highest  $\beta_{\text{leaf}}$  parameter values. After the PFT-level optimization of LPJmL-GSI herbaceous PFTs had high  $\beta_{\text{leaf}}$  parameters. The leaf albedo parameter  $\beta_{\text{leaf}}$  was sensitive in all PFTs (Figure 3.5 of the main text). The optimization resulted in many PFTs in leaf and litter albedo parameters that were close to the boundaries of the prior parameter ranges. This indicates missing environmental controls on surface albedo. The albedo routines of LPJmL need to be further improved to account for moisture-driven changes in surface albedo. Such improved albedo routines would allow a more accurate and constrained estimation of albedo parameters. Because of

these current limitations in the LPJmL albedo routines, albedo simulations in regions or time periods with low vegetation cover need to be assessed with care.

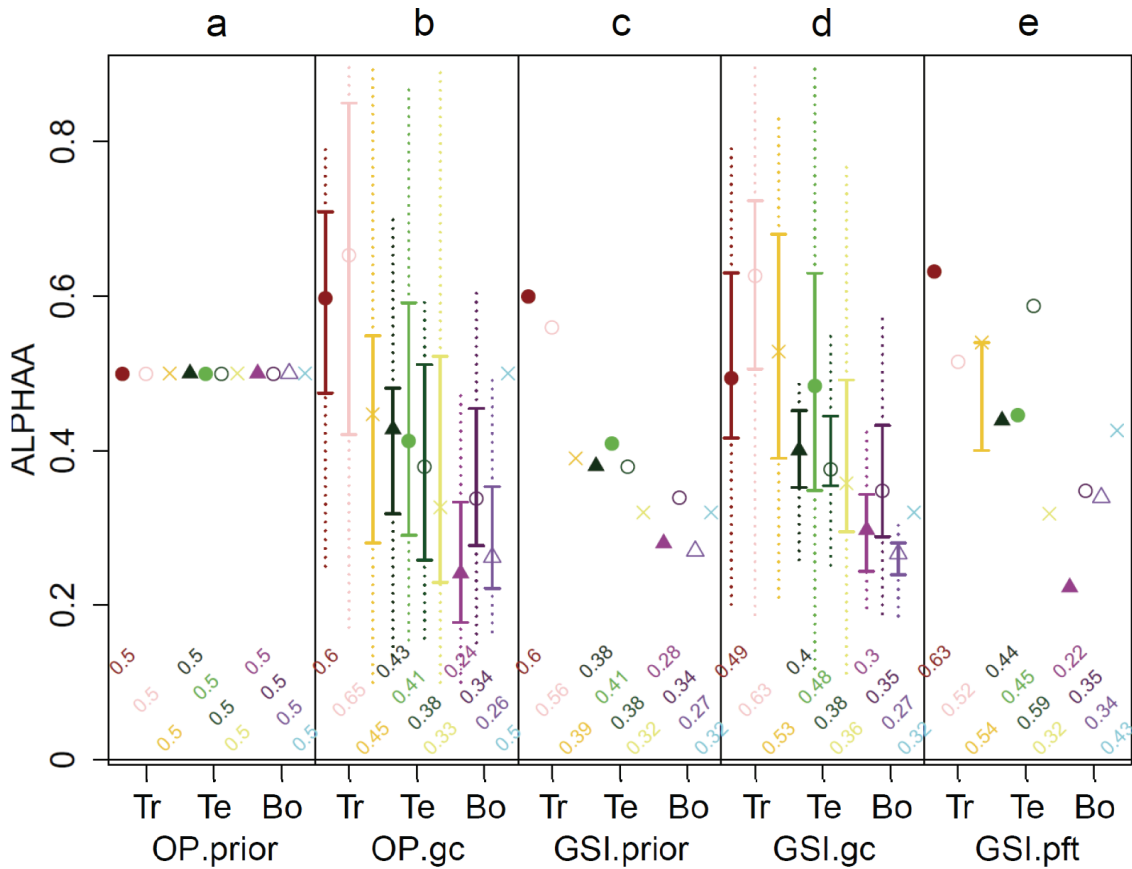


Figure 7.13: Prior and optimized values for the parameter  $\alpha$  (fraction of radiation absorbed at leaf level relative to canopy level) grouped by plant functional types and biomes. The distribution of the parameter in the optimization experiments OP.gc and GSI.gc represents the spatial variability of the parameter from different grid cell-level optimization experiments. See Figure 7.10 for a further explanation of this figure.

The light extinction coefficient  $k$  had a large spatial variability in all PFTs and in both grid cell-level optimization experiments of LPJmL-OP and LPJmL-GSI (Figure 7.15). The spatial variability was lower after grid cell-level optimization experiments of LPJmL-GSI than after grid cell-level optimization experiments of LPJmL-OP. The largest variability was found in evergreen PFTs (TrBE, TeBE, TeNE and BoNE). This result demonstrates that unique or PFT-dependent light extinction coefficient parameter values are not meaningful. Moreover, the spatial variability of the light extinction coefficient needs to be analyzed more detailed and perhaps replaced by a more advanced representation of canopy architecture.

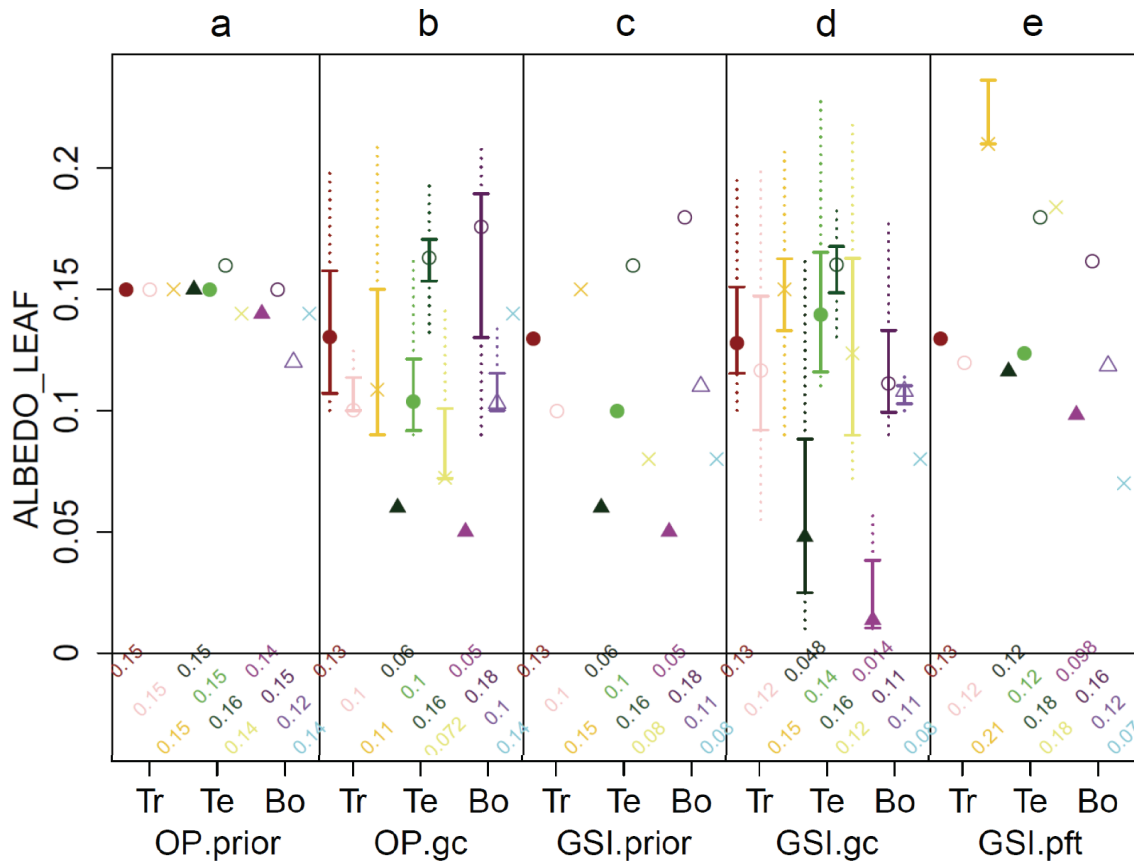


Figure 7.14: Prior and optimized values for the parameter  $\beta_{\text{leaf}}$  (leaf albedo) grouped by plant functional types and biomes.

The distribution of the parameter in the optimization experiments OP.gc and GSI.gc represents the spatial variability of the parameter from different grid cell-level optimization experiments.

See Figure 7.10 for a further explanation of this figure.

The highest values of the light extinction coefficient were found in the BoNS PFT. This was caused by an overestimation of tree mortality in years with simulated low productivity. Trees are killed in LPJmL as a result of negative net primary production which reduces FPC and results in a lower peak FAPAR in the following year. Having occurred more often in the simulated time period, it can explain why FAPAR is underestimated in some years. To remove these biases, the light extinction coefficient was optimized towards higher values in the BoNS PFT to reach FAPAR values that are closer to the observed FAPAR values after low-productivity years. However, such high values for the light extinction coefficient would overestimate tree cover and FAPAR under average conditions and when LPJmL is applied with dynamic vegetation. The approach to simulate tree mortality in LPJmL needs further improvement by, e.g., considering for example reserve carbon pools that helps the plants to endure low productivity conditions (Galvez et al., 2011).

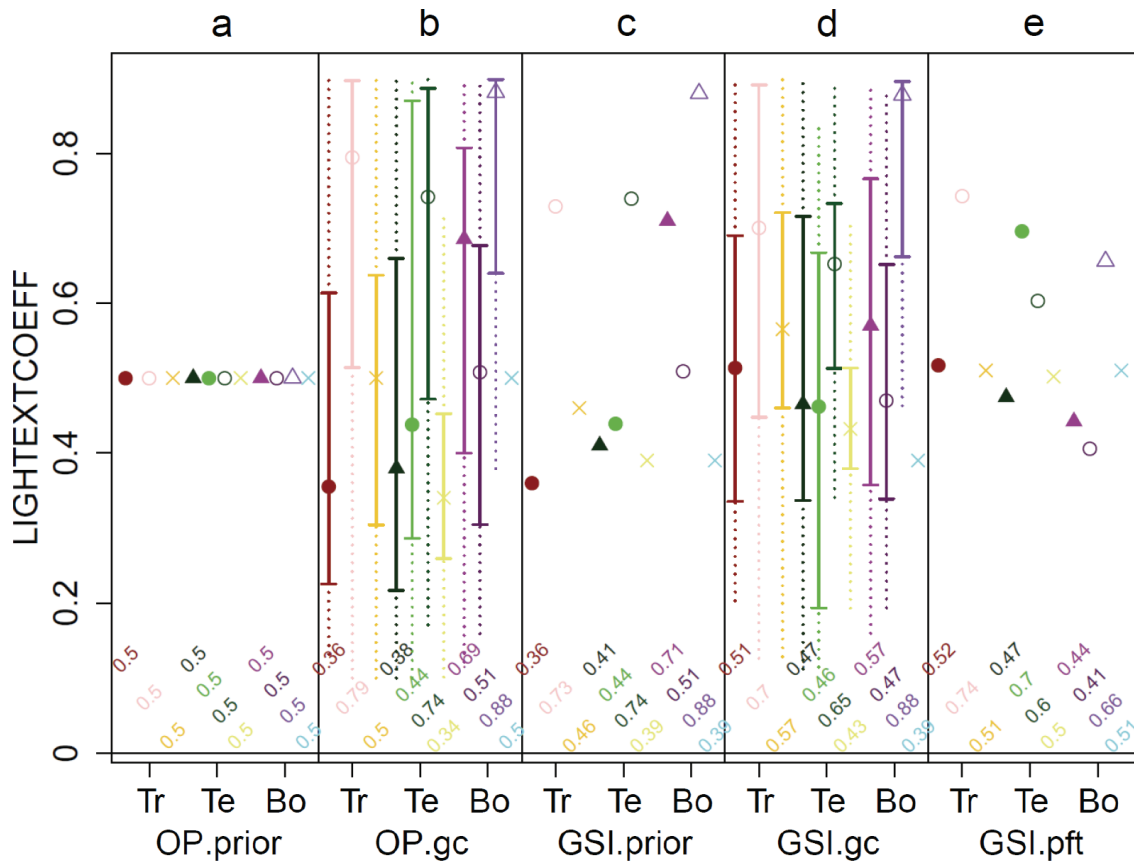


Figure 7.15: Prior and optimized values for the parameter  $k$  (light extinction coefficient) grouped by plant functional types and biomes.

The distribution of the parameter in the optimization experiments OP.gc and GSI.gc represents the spatial variability of the parameter from different grid cell-level optimization experiments. See Figure 7.10 for a further explanation of this figure.

We computed correlations between posterior parameter values for the four most important phenology parameters of LPJmL-GSI (TMIN\_BASE, LIGHT\_BASE, WATER\_BASE and TMAX\_BASE) (Figure 7.16). Most correlations were low to moderate (maximum  $r = 0.69$ ). Interestingly, the correlation between the TMIN\_BASE and WATER\_BASE parameters was low in PFTs that experience strong permafrost dynamics (BoNS  $r = 0.2$ , PoH  $r = -0.28$ ). This indicates that the water and cold temperature limiting in boreal and arctic regions were only weakly correlated. Indeed, our results showed that water availability affected phenology mostly in early spring whereas cold temperature affected phenology during the entire year in boreal and arctic regions (Figure 9 of the main text). These results emphasize the ability to disentangle effects of seasonal air temperature and soil moisture on phenology in boreal and arctic regions.



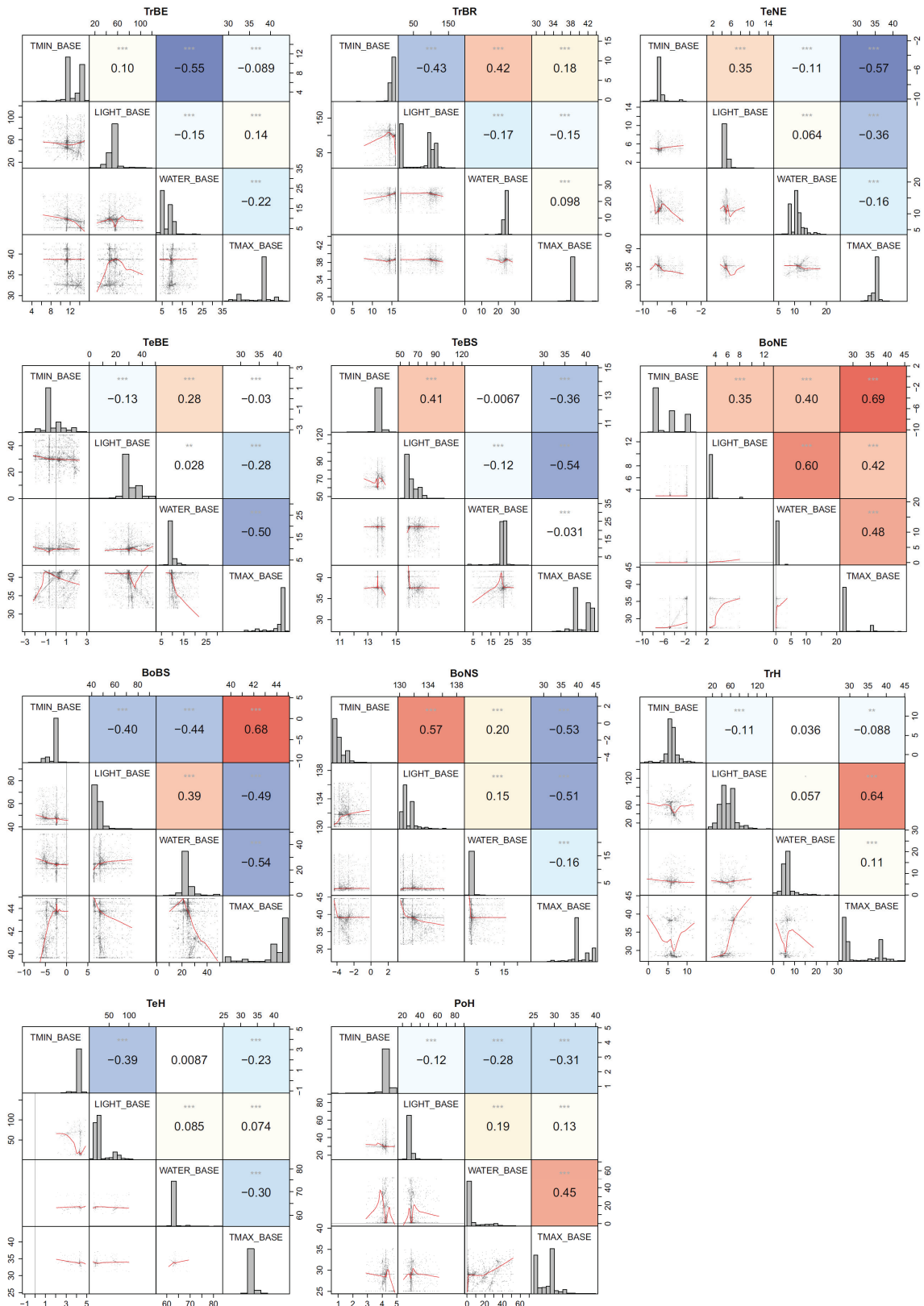


Figure 7.16: Correlations between posterior parameters for the four limiting functions for phenology in LPJmL-GSI grouped per PFT. Correlation matrices were computed based on all “best” individuals (i.e. parameters sets) from PFT-level optimization experiments (GSI.pft). “Best” individuals have an AIC difference of  $dAIC \leq 2$  in comparison to the individual with the lowest AIC, i.e. they are equally plausible. Numbers are Pearson correlation coefficients. Lines in the scatter plots are LOWESS smoothing lines (locally-weighted polynomial regression).

## 7.5 Global model evaluation

### 7.5.1 Supplementary results and discussion on carbon stocks and fluxes

LPJmL-GSI estimated global total carbon fluxes and stocks closer to data-oriented estimates than LPJmL-OP-prior and LPJmL-OP-gc (Table 7.1). All three LPJmL model versions overestimated global total GPP although LPJmL-GSI was close to the upper uncertainty estimate of the data-oriented GPP estimate. Estimates of ecosystem respiration from LPJmL were clearly larger than the data-oriented estimates. Although LPJmL simulated global total fire carbon emissions within the magnitude of independent estimates (van der Werf et al., 2010), LPJmL-OP-gc had higher and LPJmL-GSI had lower fire carbon emissions despite the use of observed burnt areas in the SPITFIRE fire module. Data-oriented estimates of global total biomass have a large uncertainty. All three version of LPJmL were within these uncertainties. LPJmL-GSI estimated global total biomass the closest to the data-oriented estimates. From Table 7.1 it is obvious that LPJmL with the model settings as in (Schaphoff et al., 2013) (i.e. without the BoNS and PoH PFTs and with simulated fire activity) resulted in global total GPP and ecosystem respiration that were even closer to the data-oriented estimates. This is mostly because LPJmL simulates larger burnt areas than seen in the observations and thus higher fire emissions but lower GPP and ecosystem respiration. Although no information about temporal variations in GPP were used in optimization experiments, the mean seasonal cycle of GPP from LPJmL-GSI and LPJmL-OP-gc agreed better with the MTE data estimate than the mean seasonal GPP cycle from LPJmL-OP-prior especially in temperate and boreal PFTs and tropical grasslands (Figure 7.18). GPP simulated by LPJmL-OP-prior increased too early and too fast in spring and decreased too late in autumn in TeNE, TeBS, BoNE, BoBS and TeH PFTs compared to the MTE estimate. These wrong dynamics improved after parameter optimization in both LPJmL-OP-gc and LPJmL-GSI. Additionally, LPJmL-GSI agreed better with the data estimate than LPJmL-OP-gc in TeNE, TeBS, TrH, PoH, TrML and TeML. These results demonstrate that the new GSI-based phenology model improved not only FAPAR seasonality but also GPP seasonality especially in temperate forests and in tropical to polar grasslands.

Table 7.6: Global total carbon fluxes and stocks from data-oriented estimates and from LPJmL simulations.

LPJmL-OP-Standard and LPJmL-GSI-Standard are LPJmL model runs with settings as in (Schaphoff et al., 2013), i.e. without the use of the BoNS and PoH PFTs and with using simulated fires instead of prescribed observed burnt areas. Data sources: 1) (Beer et al., 2010; Jung et al., 2011), 2) (van der Werf et al., 2010), 3) (Carvalhais et al., 2014; Saatchi et al., 2011; Thurner et al., 2014), 4) (Carvalhais et al., 2014).

	Gross primary production (PgC a-1)	Ecosystem respiration (PgC a-1)	Fire carbon emissions (PgC a-1)	Biomass (PgC)	Soil organic carbon (PgC)
Data estimate	124.7 <sup>1)</sup>	100-110 <sup>3)</sup>	2.0 <sup>2)</sup>	451.2 <sup>3)</sup>	2460 <sup>4)</sup>
Data lower uncertainty	110.7 <sup>1)</sup>			208.8 <sup>3)</sup>	1990 <sup>4)</sup>
Data upper uncertainty	138.3 <sup>1)</sup>			695.9 <sup>3)</sup>	2984 <sup>4)</sup>
LPJmL settings as in this study:					
LPJmL-OP-prior	161.3	150.7	1.93	674.1	2723
LPJmL-OP-gc	153.8	143.9	2.45	581.1	2503
LPJmL-GSI	145.8	141.4	1.65	546.4	2508
LPJmL settings as in Schaphoff et al. (2013):					
LPJmL-OP-Standard	138.9	125.8	3.48	597.8	2101
LPJmL-GSI-Standard	120.4	115.1	3.23	582.1	1392

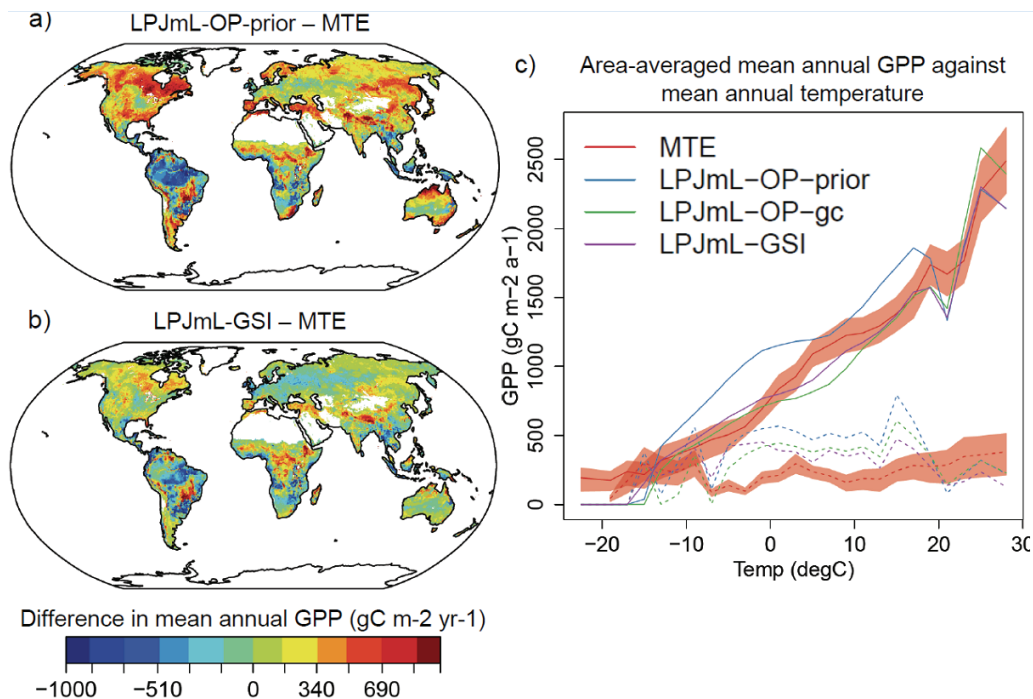


Figure 7.17: Comparison of patterns of mean annual total gross primary production from LPJmL and the data-oriented MTE estimate for the period 1982-2011.

(a) Difference in mean annual total GPP between MTE and LPJmL-OP-prior. (b) Difference in mean annual total GPP between MTE and LPJmL-GSI. (c) Global spatial-averaged gradients of mean annual GPP against mean annual temperature. Dashed lines are dry areas with mean annual P/PET < 15 and solid lines are wet areas with mean annual P/PET >= 15. The red area represents the uncertainty of the data-oriented GPP estimate expressed as the inter-quartile range of the MTE ensemble.

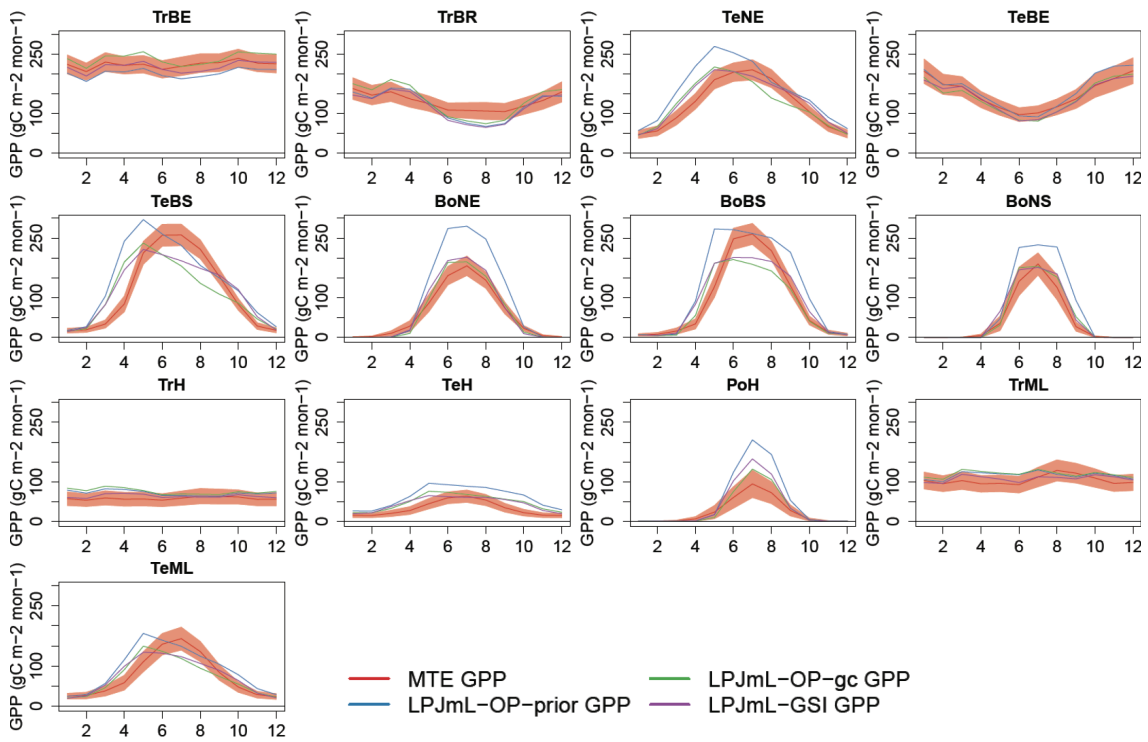


Figure 7.18: Comparison of the mean seasonal GPP cycle (averaged over 1982-2011) from MTE and LPJmL spatially averaged for regions with the same dominant PFT.

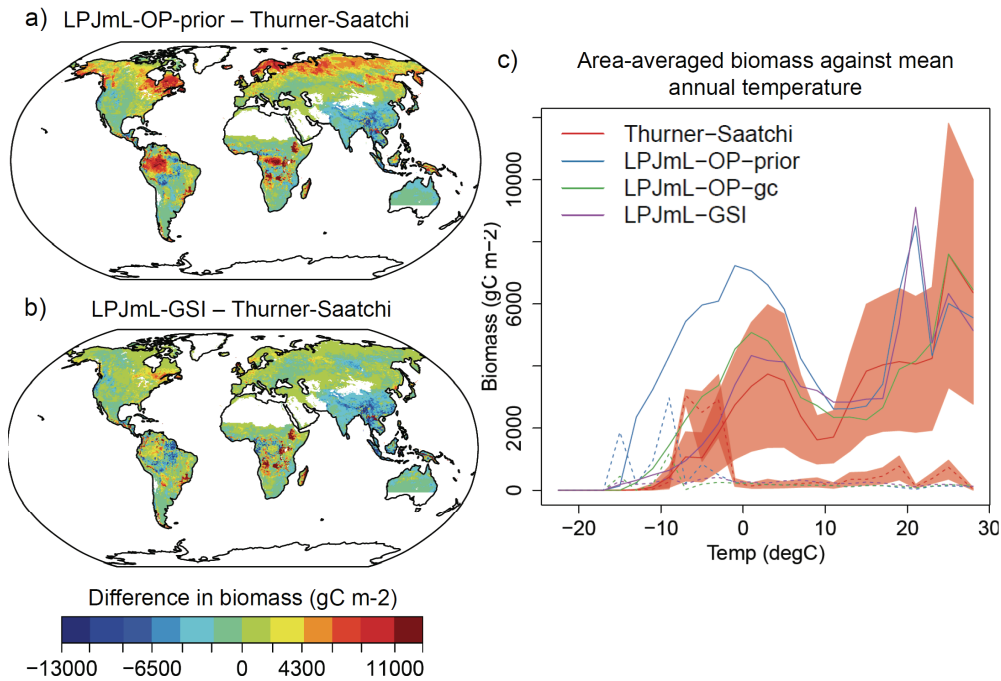


Figure 7.19: Comparison of biomass from data-oriented estimates (Thurner and Saatchi datasets) and from LPJmL (averaged 2009-2011). See Figure 7.17 for further explanations.

## 7.5.2 Supplementary figures on evapotranspiration

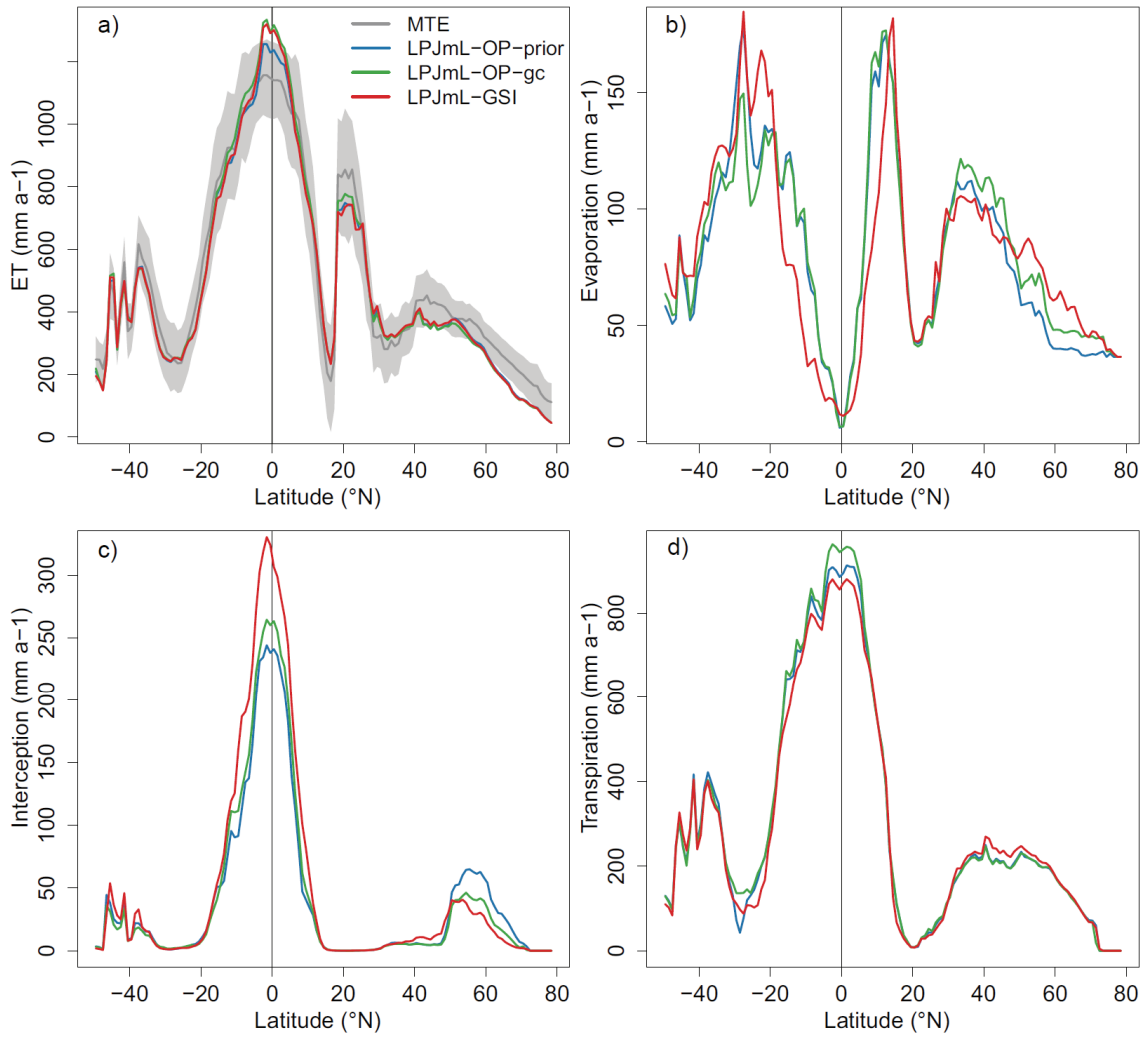


Figure 7.20: Latitudinal gradients of evapotranspiration with its components.

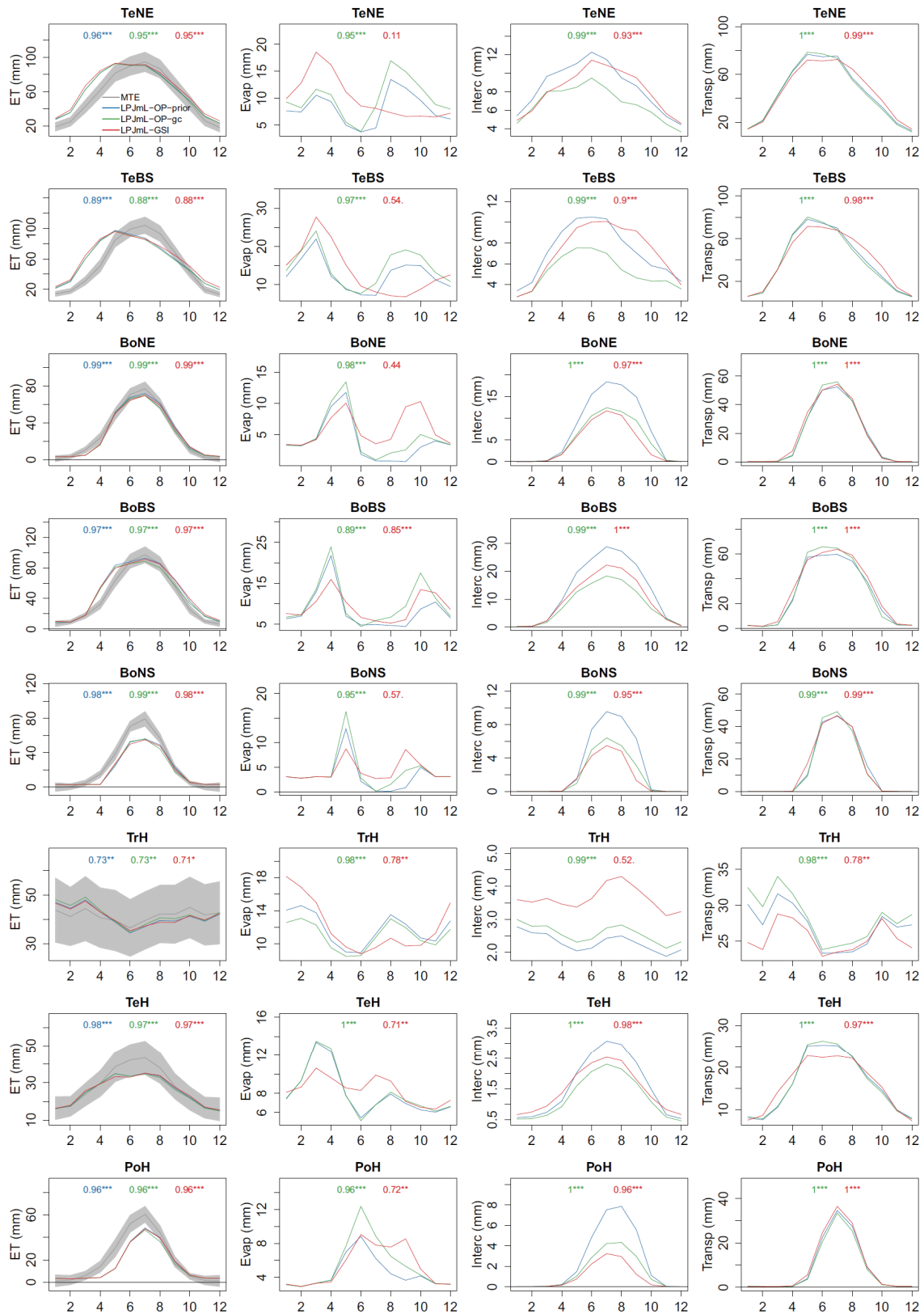


Figure 7.21: Mean seasonal cycle (1982-2011) of ET, evaporation, interception and transpiration spatially averaged for PFTs.

Numbers on top of each plot are correlation coefficients between each LPJmL model run and MTE (for ET) and between LPJmL model runs and LPJmL-OP-prior, respectively. The significance of the correlation is indicated as point symbol: \*\*\* ( $p \leq 0.001$ ), \*\* ( $p \leq 0.01$ ), \* ( $p \leq 0.05$ ), . ( $p \leq 0.1$ ).

### 7.5.3 Supplementary figures on evaluation of FAPAR

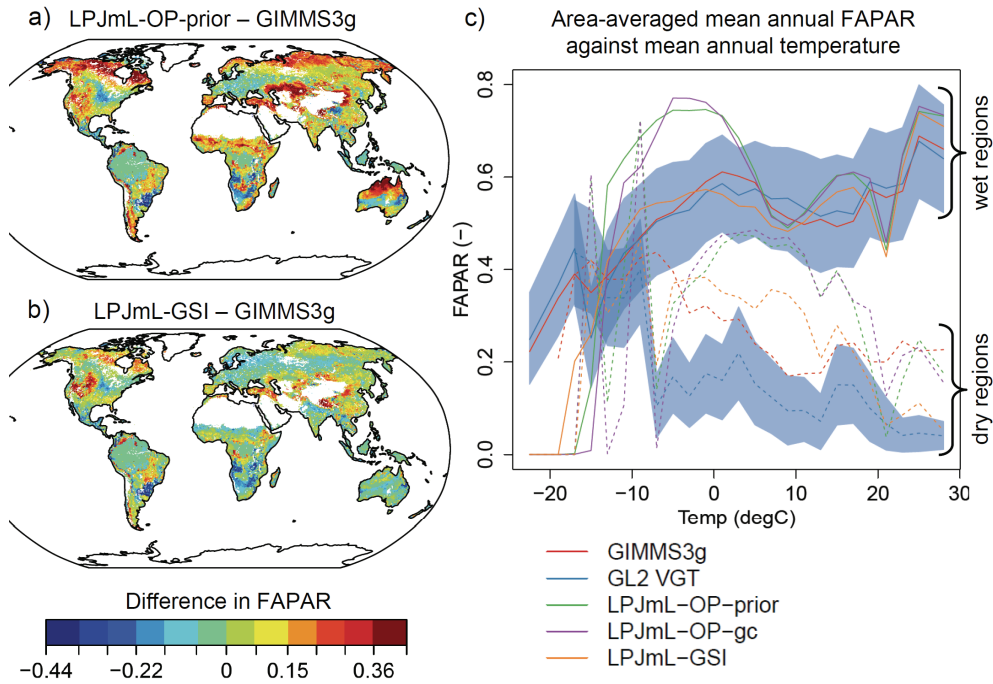


Figure 7.22: Comparison of mean annual FAPAR from LPJmL and remote sensing datasets. See Figure 7.17 for further explanations.

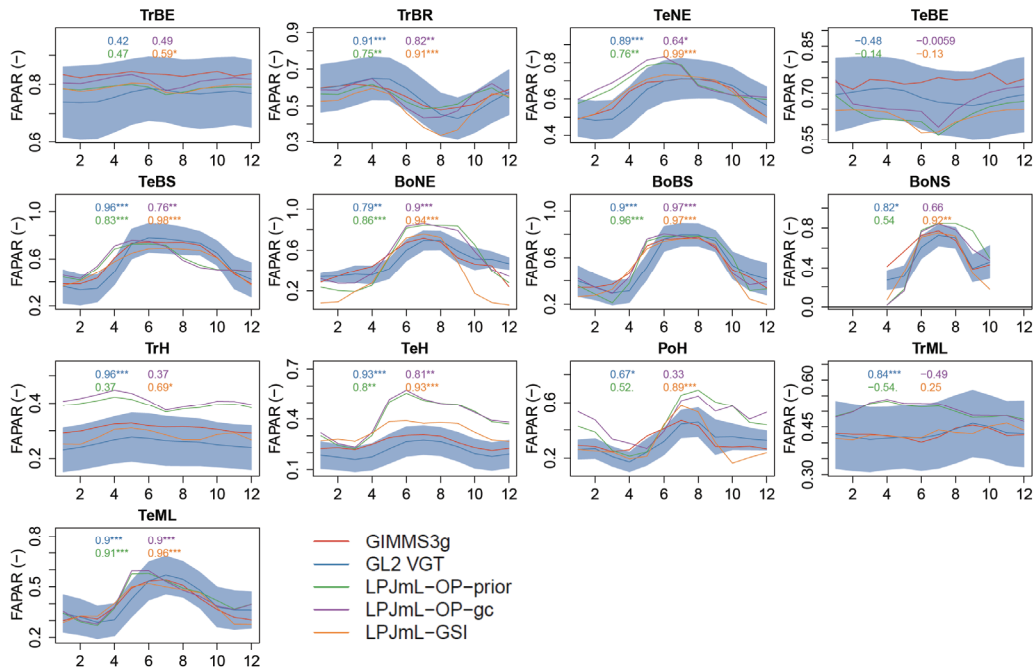


Figure 7.23: Comparison of the mean seasonal FAPAR cycle from GIMMS3g, GL2 VGT and LPJmL spatially averaged for regions with the same dominant PFT. The PFTs for which time series were averaged are shown in Figure 3. Numbers in the figures are correlation coefficients between GIMMS3g and the corresponding time series from GL2 VGT or from LPJmL simulations. The significance of the correlation is indicated as point symbol: \*\*\* ( $p \leq 0.001$ ), \*\* ( $p \leq 0.01$ ), \* ( $p \leq 0.05$ ), . ( $p \leq 0.1$ ).

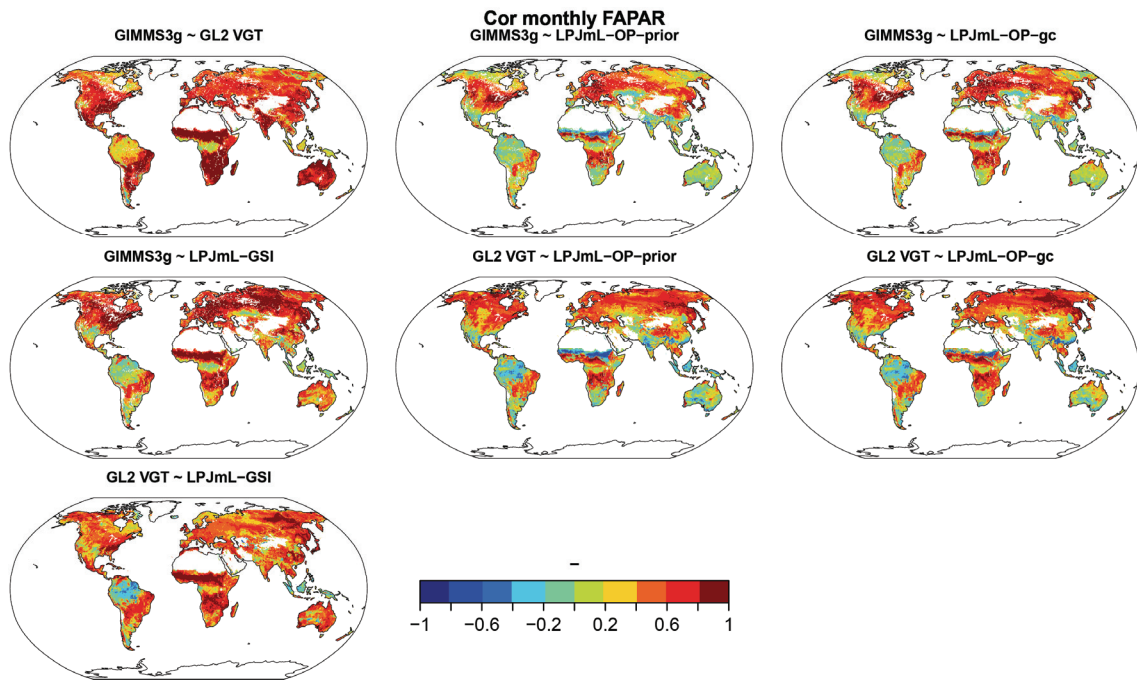


Figure 7.24: Correlation coefficients between monthly FAPAR time series from GIMMS3g, GL2 VGT datasets and LPJmL model simulations.

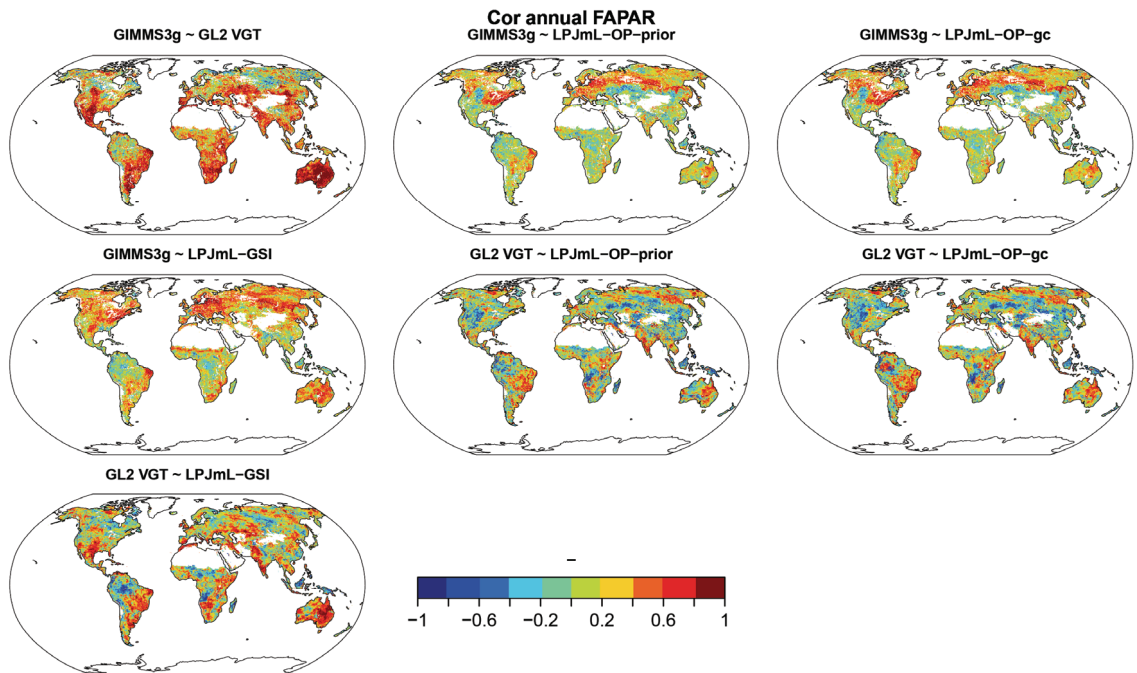


Figure 7.25: Correlation coefficients between annual FAPAR time series from GIMMS3g, GL2 VGT datasets and LPJmL model simulations. Mean annual FAPAR was averaged from monthly FAPAR values with air temperatures  $> 0^{\circ}\text{C}$ .



## Extrapolation capabilities of LPJmL-GSI

Correlation between monthly GIMMS3g and LPJmL-GSI FAPAR (1982-2011)

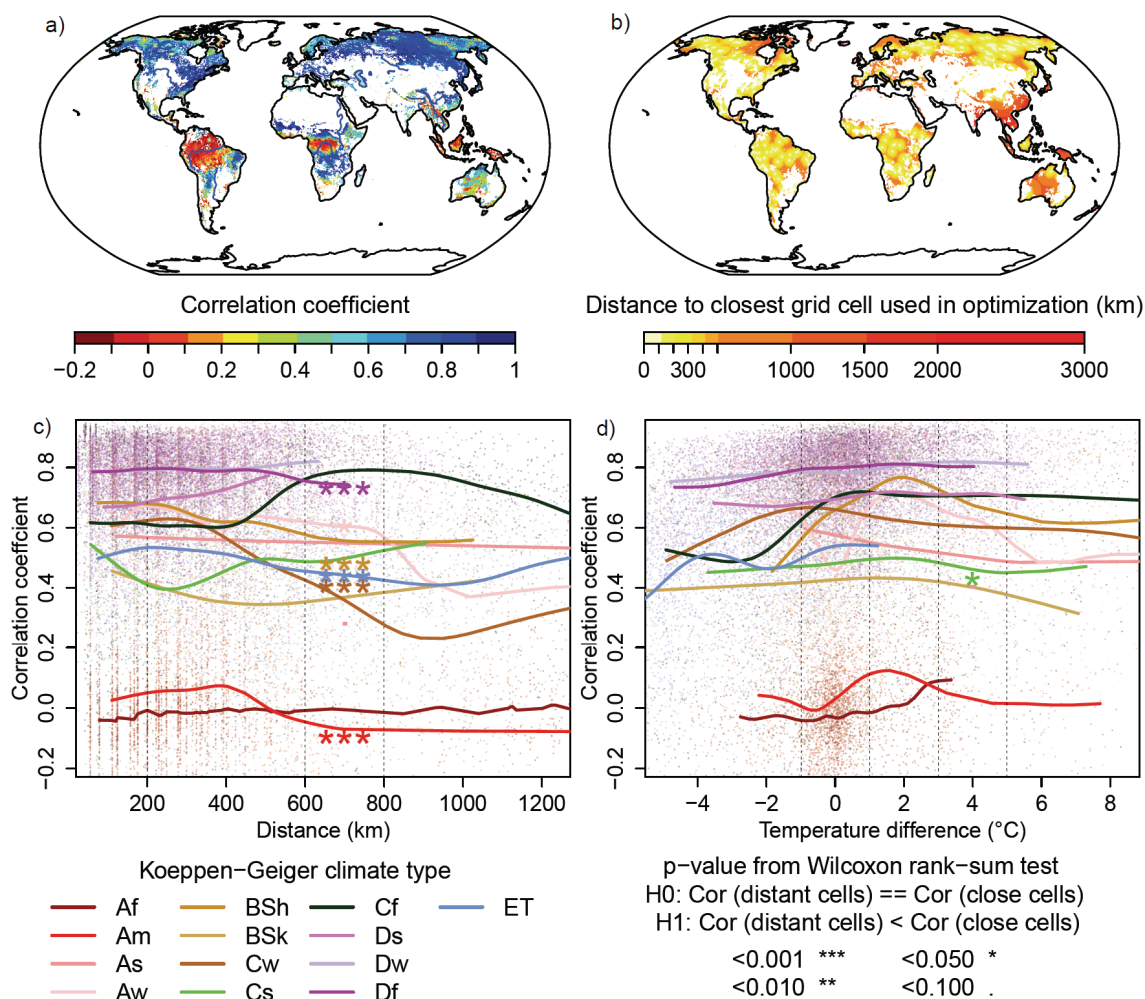


Figure 7.26: Extrapolation capabilities of LPJmL-GSI in terms of monthly FAPAR dynamics. (a) Correlation coefficient between monthly FAPAR time series from LPJmL-GSI and GIMMS3g (1982-2011). Areas without vegetation, with more than 50% agricultural use, or without data are excluded (white). (b) The map shows the distance between each  $0.5^{\circ}$  grid cell and the closest grid cell that was used in a PFT-level optimization experiment of LPJmL-GSI (GSI.pft). (c) Scatterplot between the correlation coefficient from (a) and the distance from (b) coloured by the Köppen-Geiger climate type of each grid cell. Lines are smoothing splines fitted to the quantile 0.5 of the correlation coefficient for each climate type. Star symbols indicate the p-value of a Wilcoxon rank-sum test if the correlation coefficients of distant grid cells (between 600 and 800 km, indicated by vertical dashed lines) are significant lower than of close grid cells ( $\leq 200$  km). (d) Scatterplot between the correlation coefficient from (a) and the difference in mean annual temperature between each grid cell and the corresponding closest grid cell. Star symbols indicate the p-value of a Wilcoxon rank-sum test if the correlation coefficients of warmer grid cells (between +3 and +5 $^{\circ}\text{C}$ ) are significant lower than of grid cells with similar temperature ( $\pm 1^{\circ}\text{C}$ ).

## 8 Supporting information of chapter 4

---

This chapter contains the Supporting Information of the manuscript (chapter 4):

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U. and Carvalhais, N.: Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness, *Global Change Biology*, (submitted 2014-12-16).

## 8 Supporting information of chapter 4

Table 8.1: Explanation of abbreviations.

Deriv	extreme values of the first derivative of the seasonal greenness curve to detect SOS and EOS in phenology methods
DGVM	dynamic global vegetation model
DL1	double-logistic function according to Beck et al. 2006 for time series smoothing and interpolation in phenology methods
DL2	double-logistic function according to Elmore et al. 2012 for time series smoothing and interpolation in phenology methods
eCor	effect of a factor on the correlation of a PGM time series
eMean	effect of a factor on the mean of a PGM time series
EOS	end of growing season
eTotal	total effect of a factor on a PGM time series
eTrend	effect of a factor on the trend of a PGM time series
eVar	effect of a factor on the variance of a PGM time series
FAPAR	fraction of absorbed photosynthetic active radiation
FPC	foliar projective cover
GIMMS3g	Global Inventory Modeling and Mapping Studies, 3rd generation dataset
GL2-VGT2	Geoland2 BioPar GEOV1, Vegetation 2
KGE	Kling-Gupta efficiency
LIN	linear interpolation and running median for time series smoothing and interpolation in phenology methods
LOS	length of growing season
LPJ	Lund Potsdam Jena DGVM
LPJmL	Lund Potsdam Jena managed lands DGVM
LPJmL-GSI	LPJmL model run with growing season index-based phenology module
LPJmL-OP	LPJmL model run with original phenology module
LULCC	land use and land cover change
MAU	mean autumn FAPAR
MGS	mean FAPAR of the growing season
MODIS	Moderate-Resolution Imaging Spectroradiometer
MSP	mean spring FAPAR
NDVI	normalized difference vegetation index
PEAK	annual maximum FAPAR
PFT	plant functional type
PGM	phenology and greenness metrics
PHEN	daily phenology status in LPJmL
POP	position of peak FAPAR
SOS	start of growing season
SPL	smoothing splines for time series smoothing and interpolation in phenology methods
SPOT	Satellite Pour l'Observation de la Terre
SSA	singular spectrum analysis for time series smoothing and interpolation in phenology methods
Trs	50% thresholds on the seasonal greenness curve to detect SOS and EOS in phenology methods

### 8.1 Description of phenology methods

In our approach, all phenology methods consist of three steps (Figure 4.1 of the main text): 1) Permanent gaps in each time series were filled. 2) The time series was smoothed and interpolated to daily time steps using five different methods. 3) PGMs were calculated from smoothed and daily interpolated time series using two different approaches. The in the following described methods are freely available in the R software package “greenbrown” (<http://greenbrown.r-forge.r-project.org/>).

In the first step, we identified permanent gaps as all months that have missing values in more than 20% of all years (PGAP). This implies that also an existing FAPAR observation in a month can be flagged as PGAP if missing values occur in the same month in other years. PGAP months occur usually in northern high latitude regions in winter because of snow cover or high sun zenith angles. We replaced all missing and all non-missing FAPAR values in PGAP months with the minimum FAPAR value from observations in the same PGAP months in other years or from the previous and next months, respectively. The minimum value was already used by Beck et al. (2006) to fill missing winter observation in NDVI time series.

In the second step, we used five different methods to smooth and interpolate time series to daily time steps. In method "LIN" a running median filter and linear interpolation was used. Although this method keeps as much as of the original data, it can produce some abrupt changes in the time series. In method "SPL" (Migliavacca et al., 2011) cubic splines were used to smooth (Hastie and Tibshirani, 1990) and to interpolate (Forsythe et al., 1977) the time series. In method "SSA" the time series was first linearly interpolated to daily values with added white noise on interpolated observations. Afterwards one-dimensional singular spectrum analysis (SSA) (Golyandina et al., 2001) was used to decompose the time series in components with different frequencies and finally all low frequency components (frequency  $< 1.5$ , i.e. trend to seasonal cycle) were summed to derive a smoothed time series without short-term variability. SSA was previously used to fill gaps in NDVI data (von Buttler et al., 2014) and to analyze seasonal FAPAR dynamics (Mahecha et al., 2010a). In method "DL1" a double logistic function as described in Beck et al. (2006) was fitted to the FAPAR values for each year. In method "DL2" another variant of a double logistic function was fitted to the FAPAR values for each year which better describes declining greenness during summer months (Elmore et al., 2012). Parameters of the double logistic functions in DL1 and DL2 were estimated by minimizing the sum squared error between observed and estimated FAPAR values using the BFGS optimization algorithm (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970). We excluded the estimated FAPAR values for an entire year in further analyses in case the BFGS algorithm did not converge towards an optimum value.

In the third step, PGMs were calculated in the third step using two different approaches. In the “Trs” (threshold) approach (White et al., 1997), we scaled FAPAR values of each year between 0 and 1 and computed SOS and EOS as the days when the scaled time series crosses 0.5. In the “Deriv” (derivative) approach, SOS and EOS were defined as the days when the smoothed and interpolated time series had the strongest increase and decrease, respectively (Tateishi and Ebata, 2004). Both approaches are based on the definition of SOS and EOS as the midpoints of spring greenup and autumn senescence, respectively. Although the parameters of double logistic functions (DL1 and DL2) can be directly interpreted as phenology metrics, we also calculated PGMs from these fitted functions using the “Trs” and “Deriv” approaches in order to use the same definition of SOS and EOS for all phenology methods.

PGMs as SOS, EOS and derived metrics were not calculated if a smoothed and interpolated time series had no seasonality because these metrics are meaningless in ecosystems without seasonality like evergreen tropical forests. We checked for seasonality in a time series by applying three methods to the smoothed and daily interpolated time series:

1. Periodogram: We calculated a periodogram of the time series by using fast Fourier transformation (R Core Team, 2014; Veneables and Ripley, 2002). A maximum value of the periodogram at a frequency of 1 indicates seasonality.
2. Auto-correlation function: We computed the auto-correlation function of the de-trended time series (R Core Team, 2014; Veneables and Ripley, 2002). For this we first removed a non-linear trend from the time series. The non-linear trend was estimated by using seasonal decomposition of time series by Loess (STL) (Cleveland et al., 1990). A minimum value of the auto-correlation function of the de-trended time series at a time lag of 0.5 (frequency of the time series) indicates seasonality. We assigned seasonality to the time series if the auto-correlation function had a minimum at a time lag between 0.4 and 0.6.
3. Season-trend model: We fitted two linear regression models to the time series. The first regression considers only a linear trend (trend model). The second regression considers a linear trend and seasonal factors (season-trend model). We computed the BIC (Bayesian Information Criterion) of both models. If the

BIC of the season-trend model is lower than of the trend model, the time series has probably seasonality.

We only finally assigned seasonality to a time series if all three methods indicated seasonality. If one or more of these methods indicated no seasonality, we only computed the PGMs POP and PEAK that are not depending on the estimation of SOS and EOS.

## 8.2 Supporting figures

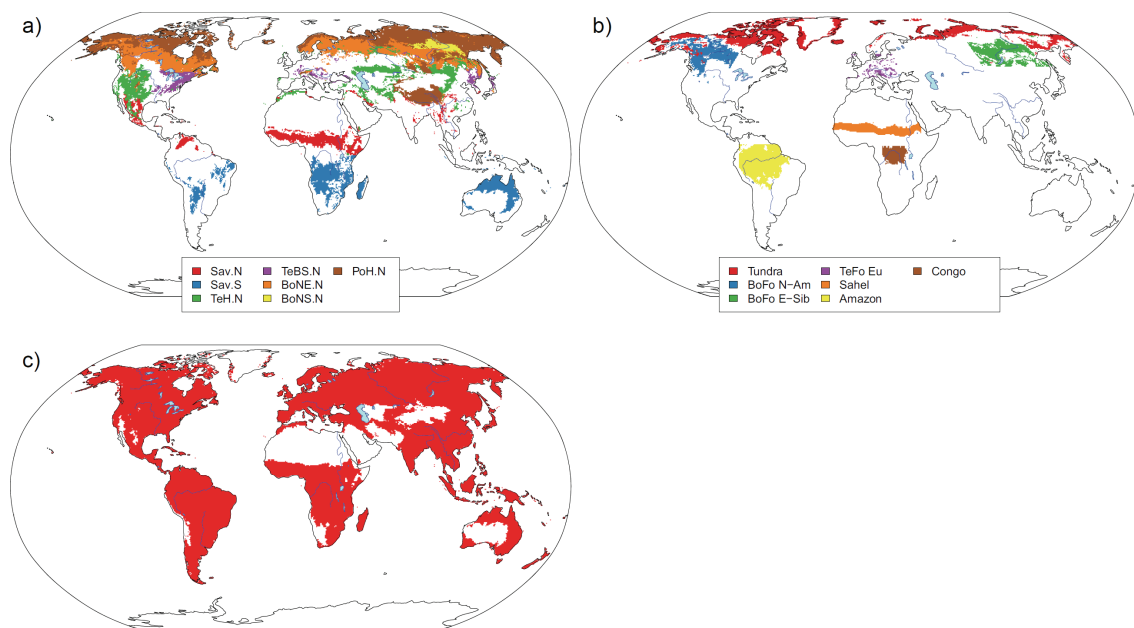


Figure 8.1: Used masks for (a) biome-aggregated, (b) region-aggregated and (c) global-aggregated results in tables and figures of the main text.

Masks for biomes and regions were derived from the PFT map as described in Forkel et al. (2014).

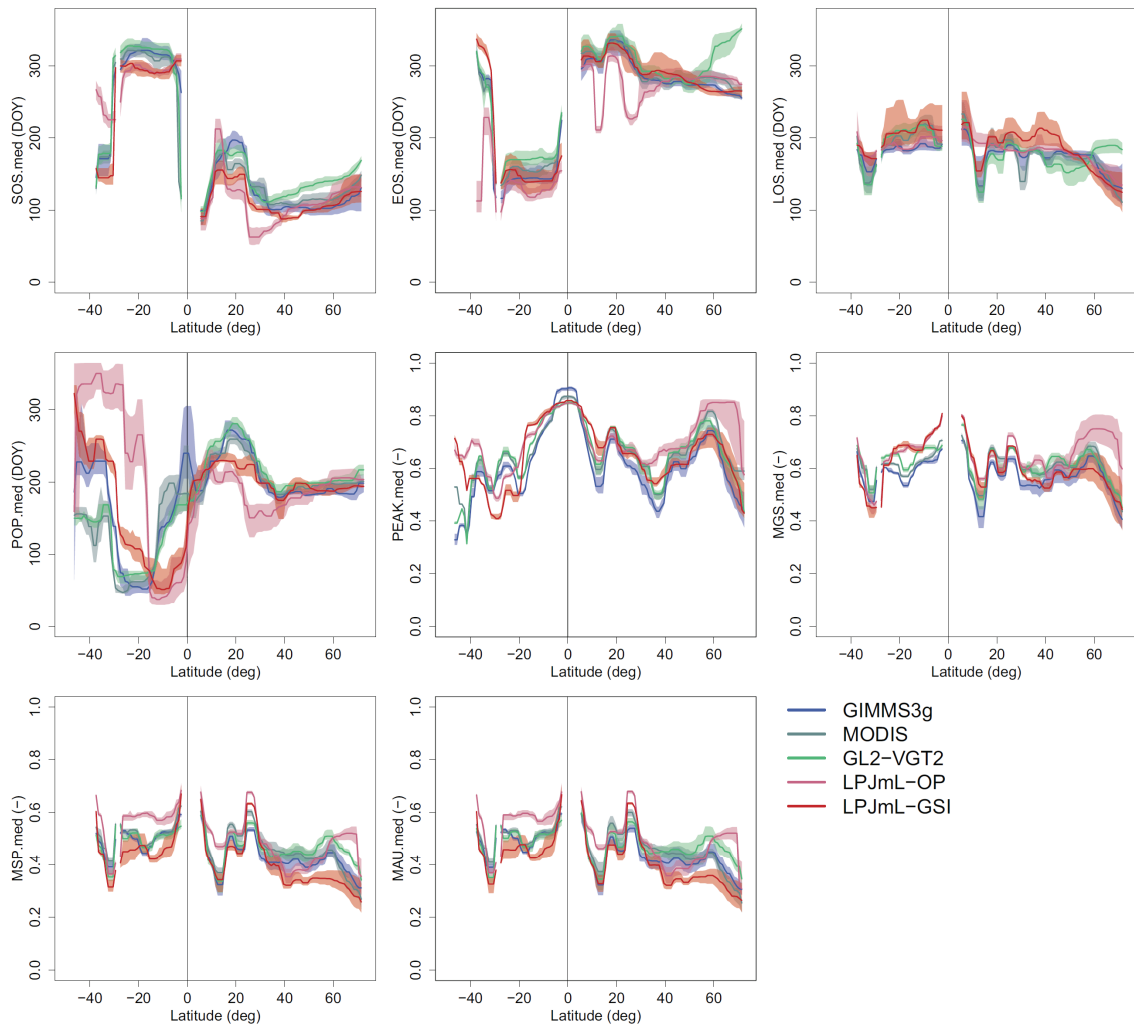


Figure 8.2: Latitudinal gradients of all PGMs. See Figure 4.2 of the main text for an explanation.

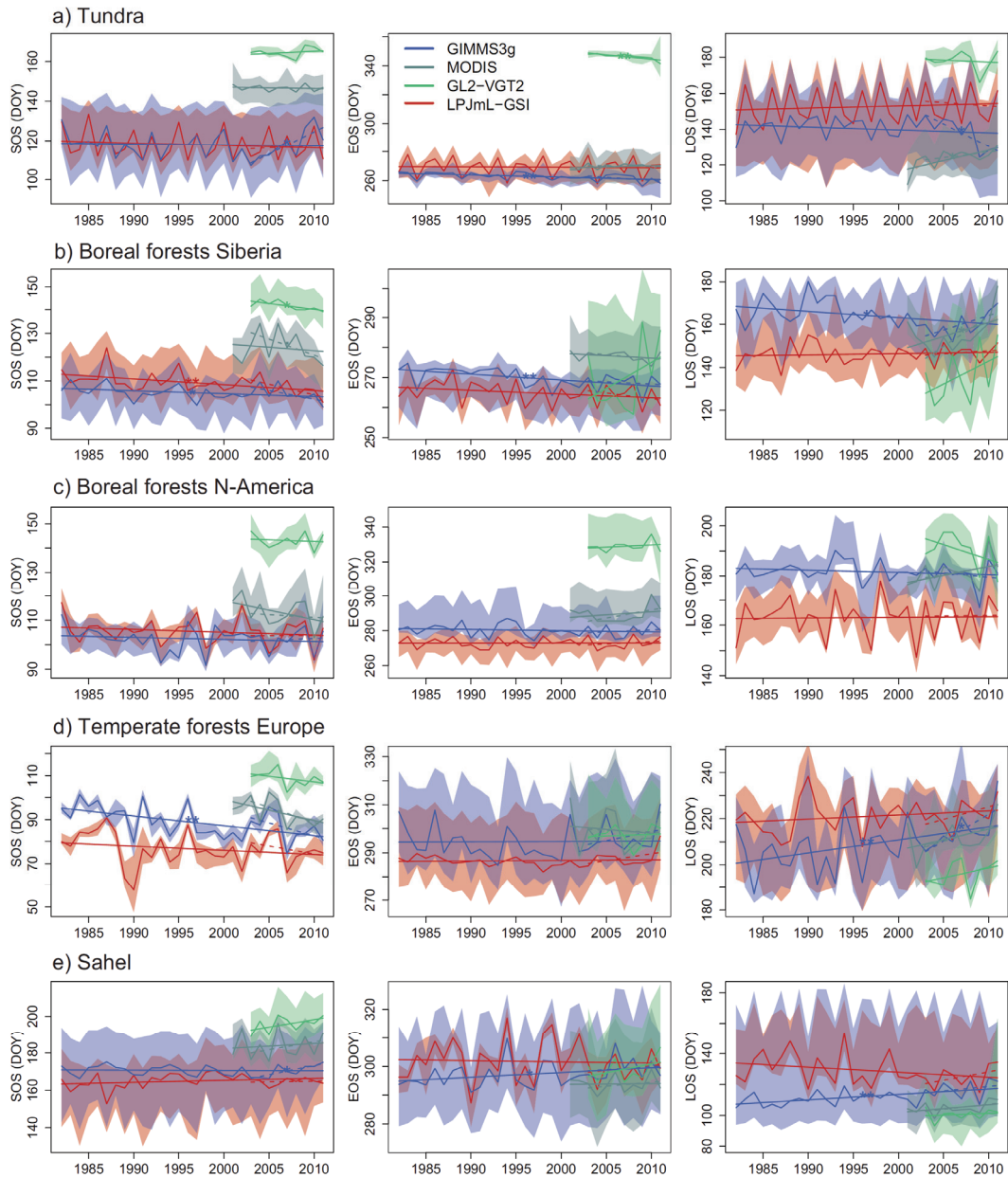


Figure 8.3: Regional-averaged time series and trends in SOS, EOS and LOS. See Figure 3 of the main text for an explanation.

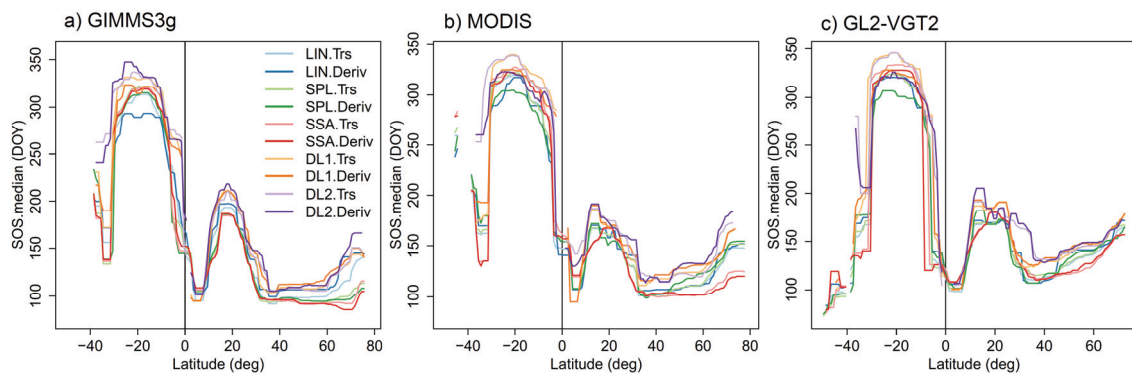


Figure 8.4: Latitudinal gradients of median annual SOS as estimated from 10 phenology methods for (a) GIMMS3g, (b) MODIS and (c) GL2-VGT2.



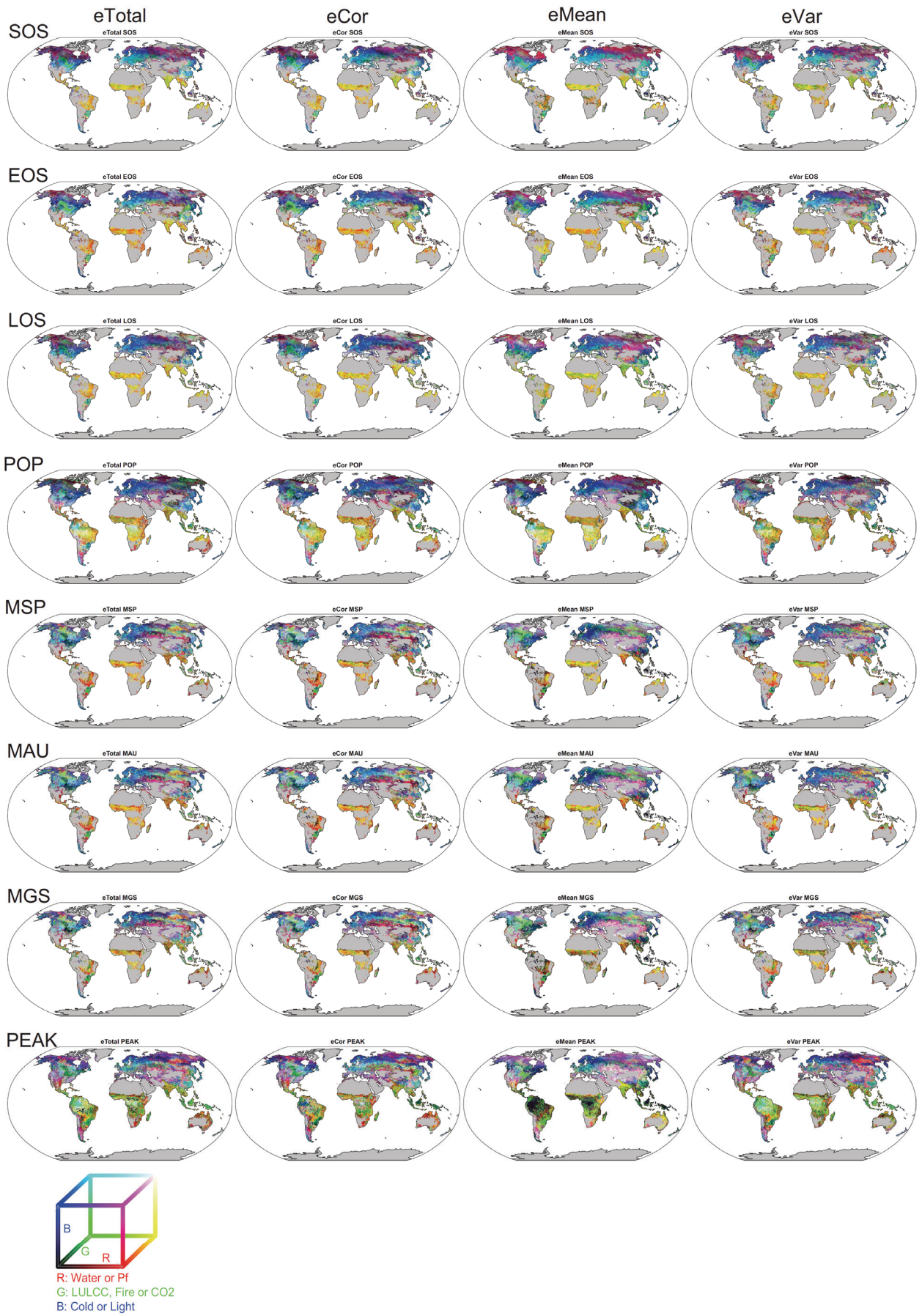


Figure 8.5: RGB composite maps of all effects on all PGMs.

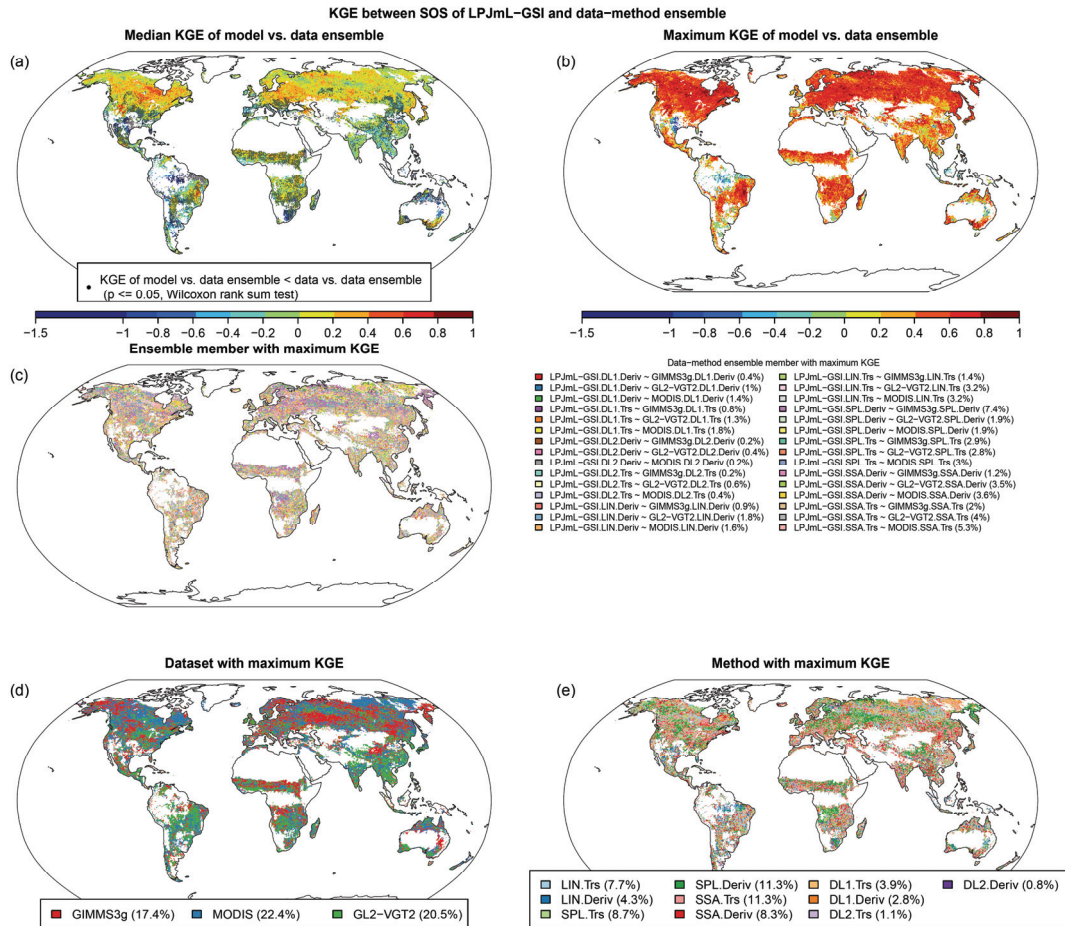


Figure 8.6: KGE between SOS from LPJmL-GSI and the data-method ensemble. (a) Ensemble median KGE. Stippling indicates grid cells for which the KGE ensemble of LPJmL-GSI is significant lower than the KGE ensemble of satellite datasets. (b) Ensemble maximum KGE. (c) Combination of satellite dataset and phenology method that resulted in the maximum KGE. (d) Satellite dataset with maximum KGE. (e) Phenology method with maximum KGE.

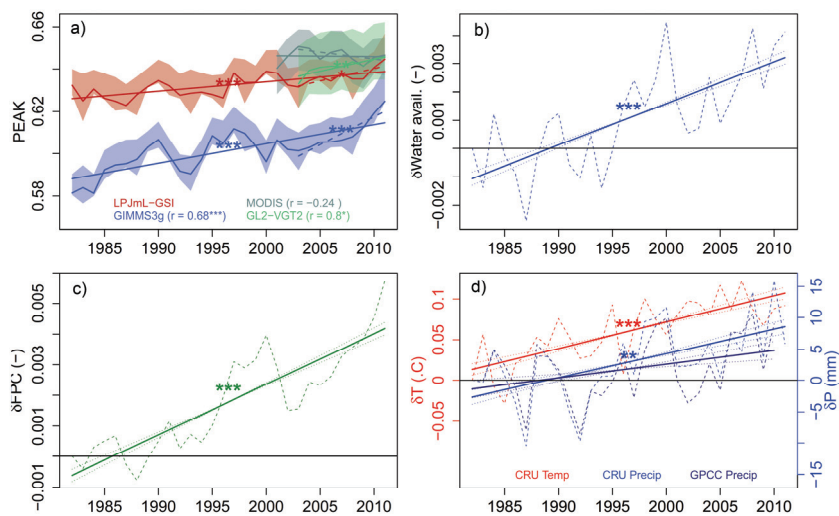


Figure 8.7: Globally spatial averaged time series and trends of (a) peak FAPAR, (b) water availability anomalies from LPJmL, (c) anomalies in total tree foliar projective cover from LPJmL and (d) air temperature and precipitation anomalies. All anomalies are relative to 1982.

## 9 Methods and extended data of chapter 5

---

This chapter contains the Methods and Extended Data of the manuscript (chapter 5):

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R. F., Heimann, M., Thonicke, K., Zaehle, S. and Reichstein, M.: Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant productivity in the northern biosphere, submitted to Nature, 2015-01-30.

## 9 Methods and extended data of chapter 5

### 9.1 Methods

#### 9.1.1 CO<sub>2</sub> time series processing and site selection

We considered using a set of 80 CO<sub>2</sub> measurement sites in our analysis but we selected only a subset of 19 sites for further analyses (Table 9.1) according to two criteria: (1) The mean seasonal CO<sub>2</sub> amplitude is larger than 5 ppm to only use sites with a pronounced seasonality. After this check no site at the southern hemisphere remained. (2) The site covers at least 20 years of observations to allow for trend analysis.

We used polynomial functions and harmonics to interpolate CO<sub>2</sub> observations to daily time series and Fast Fourier Transformation to smooth time series as implemented in a standard method (Thoning et al., 1989). We calculated the seasonal CO<sub>2</sub> amplitude as the peak-to-trough magnitude of the de-trended seasonal cycle of a calendar year. The number and temporal distribution of observations (or gaps, respectively) within a year can strongly affect the estimation of the seasonal cycle and thus the seasonal amplitude. Therefore we performed an additional “good” data coverage quality check to remove entire years from further analyses whose seasonal cycles are likely affected from outliers, few observations or an uneven temporal distribution of observations. A year with “good” data coverage had at least 12 measurements that were equally distributed among all months. To quantify how equally measurements were distributed among months, we computed the Gini coefficient  $G$  (Zeileis, 2014):

$$G = 2 \times \frac{\sum_{i=1}^N x_i \times i}{N \times \sum_{i=1}^N x_i} - \frac{N+1}{N} \quad (9.1)$$

where  $x$  is an increasing-ordered vector of the number of measurements per month and  $N$  is the number of months per year ( $N = 12$ ).  $G$  ranges between 0 (equal distribution of measurements among months) and 1 (i.e. all measurements taken in one month). We only used the interpolated and smoothed seasonal cycle of a year in further analysis if  $G \leq 0.4$ . This good data coverage quality check introduces again missing values of the length of entire years in interpolated CO<sub>2</sub> time series. We did not

use a site in further analyses if more than 50% gaps were introduced because of this quality check. These quality checks avoid drawing conclusions from CO<sub>2</sub> amplitudes of single years with a poor coverage of measurements.

### **9.1.2 Trend analysis, uncertainty estimation and latitudinal gradients**

Trends in annual time series of CO<sub>2</sub> amplitudes were computed as the slope of linear trends based on ordinary least squares regression. The significance of the trend was computed by using the Mann-Kendall trend test. We computed uncertainties in the estimated trend by adjusting time series length (McKinley et al., 2011) (Figure 9.1). This gives for each site a distribution of trend slopes. Uncertainties in trend slopes are reported as the 2.5 and 97.5%-iles of this distribution.

Latitudinal gradients of CO<sub>2</sub> amplitude trends were computed based on a linear regression between latitude and the CO<sub>2</sub> amplitude trend slope at each station. The uncertainty in the latitudinal gradient is affected by the uncertainty of trend slopes at each station and the selection of sites. Therefore, we took 1000 bootstrap samples of CO<sub>2</sub> measurement sites and randomly selected for each site a value from the site-specific distribution of CO<sub>2</sub> amplitude trend slopes. We are showing the median latitudinal gradient and the 2.5 to 97.5%-ile range.

### **9.1.3 LPJmL: simulation of terrestrial carbon fluxes, agriculture and vegetation dynamics**

We simulated terrestrial carbon fluxes and vegetation dynamics with the LPJmL (Bondeau et al., 2007; Sitch et al., 2003) (version 3.5) dynamic global vegetation model. LPJmL was used with recent model improvements for fire (Thonicke et al., 2010), permafrost (Schaphoff et al., 2013), as well as phenology (Forkel et al., 2014). Additionally, model parameters for leaf phenology and gross primary production were previously optimized against observation-based datasets of FAPAR and GPP which resulted in a better representation of global carbon stocks and fluxes (Forkel et al., 2014). In LPJmL, the net terrestrial carbon uptake (net biome productivity, NBP) is calculated from ecosystem respiration (Reco), fire C emissions (FireC), C emissions from crop harvest (H), and gross primary production (GPP):

$$\text{NBP} = (\text{Reco} + \text{FireC} + \text{H}) - \text{GPP} \quad (9.2)$$

GPP includes here the LPJmL-specific establishment flux that accounts for the amount of CO<sub>2</sub> needed to establish new seedlings when a plant functional type colonizes previously bare ground. The harvest flux accounts for crop yields whereas crop residues are respired and therefore included in Reco. Net ecosystem exchange is commonly defined as (Chapin et al., 2006; Schulze, 2006):

$$\text{NEE} = \text{Reco} - \text{GPP} \quad (9.3)$$

Reco, GPP and Estab were simulated by using the plant physiology and vegetation dynamics modules in LPJmL (Sitch et al., 2003). Fire emissions were estimated by prescribing observed burnt area time series to the LPJmL-SPITFIRE fire module (Forkel et al., 2014; Giglio et al., 2010). Emissions from land use change and agricultural harvest were directly simulated by using the agricultural routines of LPJmL (Bondeau et al., 2007) that consider land use changes and area changes of several types, irrigation, and an adaptation of sowing dates to long-term changes in climate. LPJmL was driven by observed temperature, precipitation (Harris et al., 2014), and radiation (Dee et al., 2011) time series. We analyzed LPJmL model results for the period 1970-2011.

We used observation-based estimates of GPP from the MTE (model tree ensemble) approach (Jung et al., 2011) and an estimate of NBP from the Jena Inversion system (version s81\_v3.6) (Rödenbeck et al., 2003) to evaluate LPJmL (Figure 9.5). We evaluated trends in simulated FAPAR against satellite FAPAR (GIMMS FPAR3g dataset, 1982-2011) (Zhu et al., 2013). However, the used LPJmL version was recently intensively improved and evaluated against this dataset, and against datasets of gross primary production, vegetation albedo, evapotranspiration, biomass, and tree cover (Forkel et al., 2014).

#### **9.1.4 TM3 atmospheric transport model, fossil fuel emissions, and net ocean CO<sub>2</sub> uptake**

We used the terrestrial carbon fluxes from LPJmL to simulate atmospheric CO<sub>2</sub> fractions at measurement sites with the atmospheric transport model TM3 (Heimann et al., 1989). We drive TM3 by meteorological data from the NCEP (Kalnay et al., 1996) and ERA-Interim (Dee et al., 2011) re-analysis datasets. The choice between these two

meteorological forcing datasets introduces differences in CO<sub>2</sub> amplitude trends (trends were with NCEP ca. 0.03-0.05 ppm yr<sup>-1</sup> lower than with ERA-Interim, Figure 9.3a-c). These differences have a similar magnitude like the CO<sub>2</sub> fertilization effect but are considerably smaller than the main effects of climate change or high latitude greening. Therefore we considered both meteorological datasets for model evaluation (Figure 2a, 3, 4) but used only the NCEP dataset in factorial model experiments (Figures 2b-c) because of the longer time series length. Fossil fuel emissions data was taken from the EDGAR 4.2 dataset (Olivier and Berdowski, 2001). For the air-sea net CO<sub>2</sub> fluxes we either used a data-based climatology (referred as OCLIM) (combining mean fluxes based on ocean-interior carbon data by Mikaloff Fletcher et al. (Mikaloff Fletcher et al., 2007) and seasonality by Takahashi et al. (Takahashi et al., 2009)), or a data-based estimate including inter-annual variability (referred as OIAV) by Rödenbeck et al. (Rödenbeck et al., 2014) based on surface-ocean pCO<sub>2</sub> data from the SOCAT data base (Bakker et al., 2014). Nevertheless, we did not find major differences in simulated CO<sub>2</sub> amplitude trends by using the OCLIM or OIAV datasets (Figure 9.3d-f). Therefore we used in all factorial model experiments the OCLIM dataset.

### 9.1.5 Factorial model experiments and quantification of contributions

We performed several model experiments with LPJmL+TM3 to quantify the relative role of different factors on CO<sub>2</sub> amplitude trends. The relative importance of each factor to the latitudinal gradient in the CO<sub>2</sub> amplitude trend is assessed by removing a controlling factor at a time, by maintaining the contributions of all other factors. The factorial model experiments and the change to factors are listed in Table 9.2.

We quantified the contribution of a component flux from a certain region (e.g. terrestrial NBP from boreal region) or from a process (e.g. fossil fuel, ocean, NBP, GPP, Reco) to the overall seasonal cycle. Therefore a TM3 simulation was performed for each component flux separately. The contribution to the overall seasonal cycle was quantified using the projection approach as in Graven et al. (Graven et al., 2013). Thereby the amplitude  $A$  of the seasonal CO<sub>2</sub> cycle  $A$  is defined as:

$$A = \sqrt{\sum_{m=1}^{m=12} C_m^2} \quad (9.4)$$

where  $C_m$  is the monthly total modelled CO<sub>2</sub> cycle (long-term CO<sub>2</sub> trend removed and annual mean pre-subtracted, so that the sum of  $C_m$  equals 0).  $C_m$  is the total of several component fluxes  $C_i$ :

$$C_m = \sum_{i=1}^N C_{mi} \quad (9.5)$$

The contribution (in ppm) of a component flux  $C_i$  to the total amplitude  $A$  is calculated as:

$$A_i = (1/A) \sum_{m=1}^{m=12} C_{mi} C_m \quad \text{with } A = \sum_{i=1}^n A_i \quad (9.6)$$

The percentage contribution of a component fluxes is given by:

$$x_i = A_i / A \times 100 \quad (9.7)$$

Percentage contributions add to 100 but can have positive or negative sign. Annual time series of contributions  $A_i$  are used calculate trends in contributions.



## 9.2 Extended data tables and figures

Table 9.1: Summary information about the used CO<sub>2</sub> measurement stations. Site measurements from several institutions were combined to a single station record within the time series smoothing and interpolation method.

Station	Lat (°N)	Long (°E)	Alt. (m)	Used start date	Used end date	Number of years	Observed Mean AMPCO <sub>2</sub> (ppm)	Measuring institution and reference
ALT	82.45	-62.52	210	01.01.1977	31.12.2011	35	15.6	NOAA, EC, CSIRO
AZR	38.75	-27.08	30	01.01.1981	31.12.2009	29	10.7	NOAA
BMW	32.27	-64.88	30	01.01.1990	31.12.2011	22	10.5	NOAA
BRW	71.32	-156.6	11	09.05.1971	27.12.2011	40.6	16.1	NOAA, SIO
CBA	55.2	-162.72	25	01.01.1979	31.12.2011	33	16.4	NOAA
FRD	49.88	-81.57	210	01.01.1990	31.12.2011	22	21.3	EC
GMI	13.43	144.78	5	01.01.1979	31.12.2011	33	6.3	NOAA
IZO	28.3	-16.48	2377	01.01.1992	31.12.2011	20	7.8	NOAA
KEY	25.67	-80.2	3	01.01.1973	22.12.2011	39	8.7	NOAA
LJO	32.87	-117.25	15	01.01.1979	31.12.2009	31	11.3	SIO
MHD	53.33	-9.9	10	03.06.1991	28.12.2011	20.6	14.6	NOAA
MID	28.22	-177.37	4	03.05.1985	20.12.2011	26.6	9.7	NOAA
MLO	19.53	-155.58	3397	12.01.1970	27.12.2011	42	6.7	NOAA, SIO
NWR	40.05	-105.63	3526	23.01.1970	27.12.2011	41.9	8.4	NOAA
RPB	13.17	-59.45	3	01.01.1988	29.12.2011	24	7.3	NOAA
RYO	39.03	141.83	230	01.01.1987	31.12.2011	25	14.4	JMA
SCH	47.92	7.92	1205	01.01.1972	30.12.2011	40	13.5	UBA
SHM	52.72	174.1	40	01.01.1986	25.12.2011	26	18.3	NOAA
WLG	36.27	100.92	3815	07.05.1991	29.12.2011	20.6	11.1	NOAA

### References

NOAA (Conway et al., 1994)  
SIO (Keeling et al., 2009)  
EC (Worthy, 2003)  
CSIRO (Francey et al., 2003)  
JMA (Watanabe et al., 2000)  
UBA (Levin et al., 1995)

Table 9.2: Overview of factorial model experiments with LPJmL+TM3.

Model experiment	TM3: Meteo	TM3: Ocean	LPJmL: Region or time scale	LPJmL: Meteo	LPJmL: CO <sub>2</sub>	LPJmL: Greening	LPJmL: Phenology	LPJmL: Permafrost
1a) Standard (Fig 2a, ED Fig. 4)	NCEP and ERAI	OCLIM	global	CRU	MLO	DYN	GSI	PFON
1b) Standard (all other figures)	NCEP	OCLIM	global	CRU	MLO	DYN	GSI	PFON
2) no climate change (Fig. 2b)	NCEP	OCLIM	global	CRU	MLO	DYN	GSI	PFON
3) no CO <sub>2</sub> fertilization (Fig. 2b)	NCEP	OCLIM	global	CRU	CO2FIX	DYN*	GSI	PFON
4) no high latitude greening (Fig. 2c)	NCEP	OCLIM	HL	CRU	MLO	FPCFIX	GSI	PFON
5) no high latitude climate change (Fig. 2c)	NCEP	OCLIM	HL	CRU	MLO	PRES	GSI	PFON
6) LPJmL-NCEP (ED Fig. 3a)	NCEP	OCLIM	global	CRU	MLO	DYN	GSI	PFON
7) LPJmL-NCEPfix (ED Fig. 3a)	NCEPfix	OCLIM	global	CRU	MLO	DYN	GSI	PFON
8) LPJmL-ERA1 (ED Fig. 3a)	ERA1	OCLIM	global	CRU	MLO	DYN	GSI	PFON
9) LPJmL-Mikaloff (ED Fig. 3b)	NCEP	OCLIM	global	CRU	MLO	DYN	GSI	PFON
10) LPJmL-SOCAT (ED Fig. 3b)	NCEP	OIAV	global	CRU	MLO	DYN	GSI	PFON
11) LPJmLfix-SOCAT (ED Fig. 3b)	NCEP	OIAV	MSC	CRU	MLO	DYN	GSI	PFON
12) Arctic NBP (ED Fig. 6b, ED Tab. 3a)	NCEP	OCLIM	arctic	CRU	MLO	DYN	GSI	PFON
12) Boreal NBP (ED Fig. 6b, ED Tab. 3a)	NCEP	OCLIM	boreal	CRU	MLO	DYN	GSI	PFON
12) Temperate NBP (ED Fig. 6b, ED Tab. 3a)	NCEP	OCLIM	temp.	CRU	MLO	DYN	GSI	PFON
12) Tropical NBP (ED Fig. 6b, ED Tab. 3a)	NCEP	OCLIM	tropical	CRU	MLO	DYN	GSI	PFON
13) Cropland NBP (ED Fig. 6b, ED Tab. 3a)	NCEP	OCLIM	cropland	CRU	MLO	DYN	GSI	PFON
14) LPJmL-oldPhen (ED Fig. 6c)	NCEP	OCLIM	global	CRU	MLO	DYN	OP	PFON
15) LPJmL-noPF (ED Fig. 6c)	NCEP	OCLIM	global	CRU	MLO	DYN	OP	PFOFF
<b>Explanation of factors</b>								
TM3: Meteo	To test the influence of meteorological forcing dataset for TM3: NCEP: NCEP reanalysis (1970-2011) ERA1: ERA-Interim reanalysis (1980-2011) NCEPfix: meteorology from NCEP of the year 1995 was used for all years							
TM3: Ocean	To test the effect of ocean CO <sub>2</sub> uptake: OCLIM (Mikaloff): dataset from Mikaloff Fletcher et al.41, without inter-annual variability and trends OIAV (SOCAT): dataset from Rödenbeck et al.22 better accounts for inter-annual variability							
LPJmL: region or time scale	To test the contribution of NBP from different regions, or temporal scales: global: all factors were applied globally HL: the factorial change was only applied in high latitude ecosystems (Köppen-Geiger D and E climate zones, > 41°N), in the rest of the world simulation from the standard LPJmL model run were used MSC: the mean seasonal cycle of NBP from the LPJmL standard model run was used for all years arctic/boreal/temperate/tropical/cropland: NBP from the LPJmL standard model run was cutout for							

	the specific region and rest of the world was set to NBP = 0
LPJmL: Meteo	To test the effect of climate trends and changing climate variability in LPJmL: CRU: temperature, precipitation from CRU TS3.2, radiation from ERA-Interim CRUFIX: CRU climate data from the years 1965-1975 was randomly repeated in the period 1970-2011
LPJmL: CO2	To test the effect of CO2 fertilization in LPJmL: MLO: LPJmL is forced with mean annual CO2 time series from MLO. CO2FIX: LPJmL is forced with a constant mean annual CO2 after 1970 (325.68 ppm)
LPJmL: greening	To test the effect of vegetation dynamics ("greening"): DYN: simulated vegetation dynamics DYN*: simulated vegetation dynamics but transient trends change because of fixing another factor in LPJmL (CRUFIX or CO2FIX) FPCFIX: foliar projective cover (fractional area coverage of plant functional types) was fixed to constant values after 1970 in high-latitude regions at the simulated value of 1970 PRES: foliar projective cover of plant functional types as simulated in the standard model run was prescribed to LPJmL to keep vegetation dynamics as in standard model run while fixing another factor which would usually change vegetation dynamics
LPJmL: phenology	To test the effect of a new phenology module and optimized model parameters in LPJmL: GSI: improved phenology scheme and optimized model parameters (Forkel et al., 2014) OP: original phenology scheme and model parameters (Sitch et al., 2003)
LPJmL: permafrost	To test the effect of the simulation of soil thermal dynamics (permafrost): PFON: simulation of permafrost PFOFF: simulation of permafrost dynamics is enabled

Table 9.3: Changes in the contribution to the seasonal amplitude of CO<sub>2</sub> separated (a) by regions and (b) by terrestrial gross fluxes.

Shown are the average total amplitude and the fractional contributions in the 1970s (1970-1979) and in the 2000s (2000-2011) as well as the trend in the contribution (% yr<sup>-1</sup>) over the full time series length at each site. Trend p-values of the Mann-Kendall test are indicated by symbols: \*\*\* p ≤ 0.001, \*\* p ≤ 0.01, \* p ≤ 0.05, . p > 0.05. Percentage contributions may not equal 100% and the sum of contributing trends may not equal the total trend because of rounding. Stations are ordered by latitude from north to south. Please refer to equations 4-7 for the calculations of these metrics.

(a) Station and period	Ampli- tude (A)	Contribution of regional fluxes to total amplitude (x <sub>i</sub> and trend in A <sub>i</sub> )					
		Fossil fuel + ocean	Arctic NBP	Boreal NBP	Temperate NBP	Tropical NBP	Cropland NBP
ALT 1970s	16.2 ppm	5 %	21 %	59 %	13 %	-7 %	9 %
ALT 2000s	18.2 ppm	2 %	22 %	56 %	14 %	-6 %	11 %
ALT trend	0.64 %/yr ***	-0.13 %/yr ***	0.23 %/yr ***	0.28 %/yr **	0.1 %/yr .	0.03 %/yr .	0.14 %/yr *
BRW 1970s	17.1 ppm	5 %	26 %	54 %	12 %	-5 %	8 %
BRW 2000s	20.3 ppm	2 %	32 %	52 %	11 %	-5 %	9 %
BRW trend	0.63 %/yr ***	-0.09 %/yr ***	0.35 %/yr ***	0.26 %/yr ***	0.04 %/yr .	-0.02 %/yr .	0.1 %/yr **
CBA 1970s	15.1 ppm	7 %	14 %	57 %	20 %	-5 %	8 %
CBA 2000s	17.6 ppm	7 %	13 %	57 %	19 %	-5 %	9 %
CBA trend	0.63 %/yr **	-0.07 %/yr .	0.15 %/yr *	0.35 %/yr *	0.06 %/yr .	0.02 %/yr .	0.13 %/yr **
SCH 1970s	19.8 ppm	1 %	5 %	35 %	46 %	-7 %	19 %
SCH 2000s	21.9 ppm	0 %	5 %	35 %	44 %	-6 %	21 %
SCH trend	0.37 %/yr *	-0.03 %/yr .	0.04 %/yr ***	0.11 %/yr **	0.11 %/yr .	0.01 %/yr .	0.14 %/yr **
NWR 1970s	5.3 ppm	-16 %	10 %	59 %	27 %	17 %	3 %
NWR 2000s	6 ppm	-11 %	11 %	58 %	22 %	12 %	8 %
NWR trend	1.15 %/yr	-0.21 %/yr .	0.2 %/yr **	0.58 %/yr .	0.13 %/yr .	0.09 %/yr .	0.36 %/yr .
LJO 1970s	5.2 ppm	-2 %	0 %	-12 %	21 %	75 %	19 %
LJO 2000s	4.7 ppm	2 %	14 %	68 %	0 %	2 %	13 %
LJO trend	-0.01 %/yr	0.26 %/yr .	0.33 %/yr **	1.61 %/yr **	-0.33 %/yr .	-1.7 %/yr **	-0.17 %/yr .
KEY 1970s	11.1 ppm	-7 %	4 %	26 %	12 %	48 %	17 %
KEY 2000s	13.2 ppm	-1 %	4 %	24 %	11 %	48 %	14 %
KEY trend	0.98 %/yr ***	0.2 %/yr *	0.03 %/yr *	0.11 %/yr *	0.08 %/yr .	0.52 %/yr *	0.03 %/yr .
MLO 1970s	7.1 ppm	-10 %	7 %	40 %	23 %	25 %	15 %
MLO 2000s	7.2 ppm	-9 %	9 %	46 %	22 %	18 %	14 %
MLO trend	0.28 %/yr	-0.01 %/yr .	0.12 %/yr ***	0.33 %/yr ***	0.03 %/yr .	-0.19 %/yr .	0.01 %/yr .
GMI 1970s	4.8 ppm	3 %	4 %	27 %	21 %	34 %	11 %
GMI 2000s	6.3 ppm	9 %	5 %	33 %	18 %	23 %	12 %
GMI trend	1.37 %/yr **	0.32 %/yr **	0.1 %/yr **	0.45 %/yr ***	0.11 %/yr .	0.3 %/yr .	0.1 %/yr .
<b>Averages over high latitude (HL &gt; 45°N) and low latitude (LL &lt; 45°N) sites:</b>							
HL 1970s	17 ppm	4 %	17 %	51 %	23 %	-6 %	11 %
HL 2000s	18.9 ppm	3 %	16 %	54 %	20 %	-5 %	12 %
HL trend	0.71 %/yr **	-0.1 %/yr .	0.18 %/yr *	0.36 %/yr *	0.08 %/yr .	0.05 %/yr .	0.14 %/yr .
LL 1970s	6.7 ppm	-6 %	5 %	28 %	21 %	40 %	13 %
LL 2000s	8.7 ppm	-5 %	9 %	47 %	18 %	14 %	17 %
LL trend	0.97 %/yr	0 %/yr .	0.14 %/yr .	0.46 %/yr .	0.08 %/yr .	0.12 %/yr .	0.17 %/yr .

(Table 9.3 continues on next page)

(Table 9.3 continued)

(b) Station and period	Ampli- tude (A)	Contribution of flux types to total amplitude ( $x_i$ and trend in $A_i$ )					
		Fossil fuel	Ocean	Land: GPP	Land: Reco	Land: Fire	Land: Harvest
ALT 1970s	16.9 ppm	4%	1%	261%	-165%	-1%	1%
ALT 2000s	18.6 ppm	2%	1%	239%	-141%	-1%	1%
ALT trend	0.58 %/yr ***	-0.11 %/yr ***	-0.02 %/yr .	1.7 %/yr .	-1 %/yr .	0 %/yr .	0 %/yr .
BRW 1970s	17.8 ppm	5%	0%	279%	-183%	-2%	1%
BRW 2000s	20.8 ppm	1%	1%	235%	-136%	-2%	1%
BRW trend	0.6 %/yr ***	-0.09 %/yr ***	0 %/yr .	1.32 %/yr .	-0.62 %/yr .	-0.01 %/yr .	0 %/yr .
CBA 1970s	15.7 ppm	0%	6%	285%	-191%	-1%	1%
CBA 2000s	18.2 ppm	1%	6%	240%	-146%	-1%	1%
CBA trend	0.66 %/yr ***	-0.08 %/yr *	0.02 %/yr .	1.66 %/yr .	-0.93 %/yr .	-0.01 %/yr .	0.01 %/yr .
SCH 1970s	20.3 ppm	1%	0%	177%	-78%	0%	1%
SCH 2000s	22.3 ppm	0%	0%	180%	-80%	0%	0%
SCH trend	0.37 %/yr *	-0.04 %/yr .	0 %/yr .	0.92 %/yr *	-0.51 %/yr .	0.01 %/yr .	-0.01 %/yr .
NWR 1970s	5.3 ppm	-4%	-10%	819%	-704%	-1%	0%
NWR 2000s	6.1 ppm	-3%	-8%	593%	-482%	-1%	0%
NWR trend	1.22 %/yr	-0.22 %/yr .	0.01 %/yr .	-1.58 %/yr .	3 %/yr .	0.02 %/yr .	-0.02 %/yr .
LJO 1970s	4.9 ppm	-7%	3%	332%	-228%	2%	-1%
LJO 2000s	4.4 ppm	16%	-9%	455%	-360%	-2%	1%
LJO trend	-0.55 %/yr	0.58 %/yr **	-0.13 %/yr .	-6.92 %/yr .	5.92 %/yr .	-0.04 %/yr .	0.04 %/yr .
KEY 1970s	11.9 ppm	-1%	-5%	334%	-229%	1%	0%
KEY 2000s	14.6 ppm	4%	-5%	244%	-145%	1%	0%
KEY trend	1.05 %/yr ***	0.2 %/yr **	-0.02 %/yr .	0.19 %/yr .	0.64 %/yr .	0.01 %/yr .	0.02 %/yr .
MLO 1970s	6.4 ppm	-2%	-9%	514%	-405%	1%	1%
MLO 2000s	7.5 ppm	-1%	-7%	412%	-306%	1%	1%
MLO trend	0.88 %/yr **	-0.03 %/yr .	0.03 %/yr *	-0.71 %/yr .	1.6 %/yr *	0.01 %/yr .	-0.01 %/yr .
GMI 1970s	4.9 ppm	15%	-12%	348%	-260%	3%	5%
GMI 2000s	6.2 ppm	19%	-10%	294%	-209%	2%	4%
GMI trend	1.41 %/yr **	0.43 %/yr **	-0.06 %/yr .	0.81 %/yr .	0.19 %/yr .	0.02 %/yr .	0.03 %/yr .
<b>Averages over high latitude (HL &gt; 45°N) and low latitude (LL &lt; 45°N) sites:</b>							
HL 1970s	17.6 ppm	2%	2%	250%	-154%	-1%	1%
HL 2000s	19.4 ppm	1%	2%	234%	-137%	-1%	1%
HL trend	0.71 %/yr **	-0.11 %/yr .	0 %/yr .	2.31 %/yr .	-1.5 %/yr .	0 %/yr .	0 %/yr .
LL 1970s	6.7 ppm	0%	-7%	469%	-365%	1%	1%
LL 2000s	8.9 ppm	2%	-6%	399%	-295%	0%	1%
LL trend	0.99 %/yr	0.05 %/yr .	-0.03 %/yr .	0.55 %/yr .	0.43 %/yr .	-0.01 %/yr .	0 %/yr .

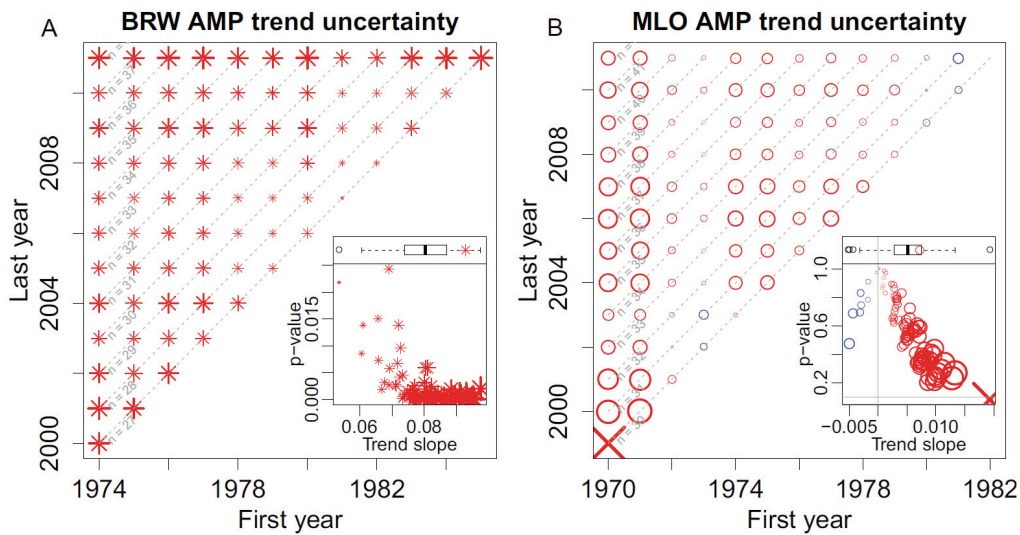


Figure 9.1: Uncertainty in observed CO<sub>2</sub> amplitude trends at (a) BRW and (b) MLO depending on the length of the time series.

The symbols indicate the trend direction (red = positive, blue = negative), the trend slope (symbol size), and the significance (symbol type). The inset scatter plot shows the relationship between trend slope against p-value (Mann-Kendall trend test) with the corresponding scaling of symbols according to trend slope. Boxplots show the distribution of trend slopes. The symbol on top of a boxplot indicates the trend of the full time series length. While for BRW the trend is significantly positive for all period lengths, for MLO some combinations of years can result in a negative (but non-significant) trend which demonstrates the need to assess the uncertainty of trends depending on the observational window.

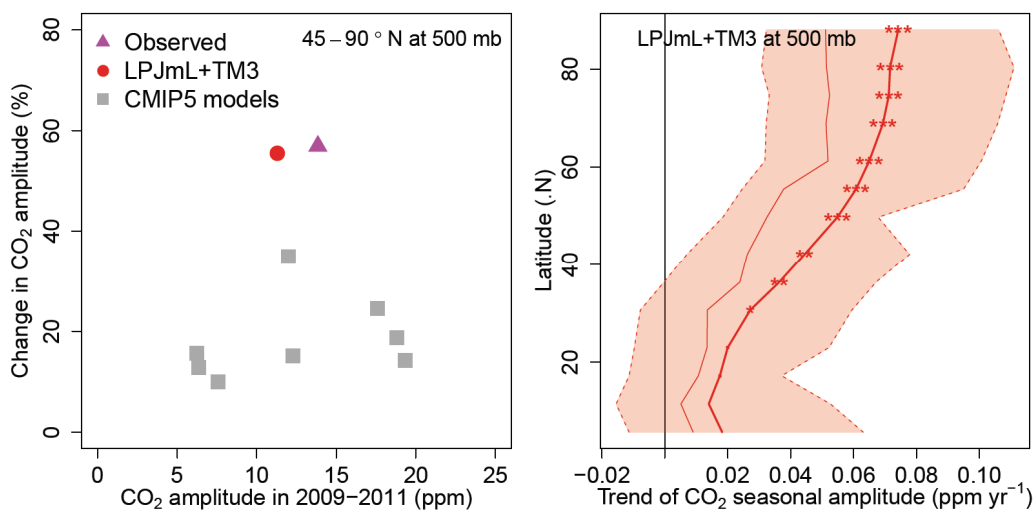


Figure 9.2: Changes in CO<sub>2</sub> amplitude at 500 mb.

(a) Change in the amplitude of the seasonal cycle of CO<sub>2</sub> versus amplitude of the seasonal cycle for 2009 to 2011 at 500 mb, averaged over 45° to 90°N. The design of the figure and results for observations and CMIP5 models were directly taken from Fig. 4 in Graven et al. (Graven et al., 2013). Change in CO<sub>2</sub> amplitude is relative to 1958-1961 for observations and CMIP5 models, and relative to 1970-1973 for LPJmL+TM3. (b) Latitudinal gradient of the CO<sub>2</sub> amplitude trend at 500 mb as simulated with LPJmL+TM3. The bold line with significance flags is the trend using the full time series length whereas thin lines and bands represent the 50%-ile, and 2.5% to 97.5%-iles of the trend uncertainty estimate, respectively.

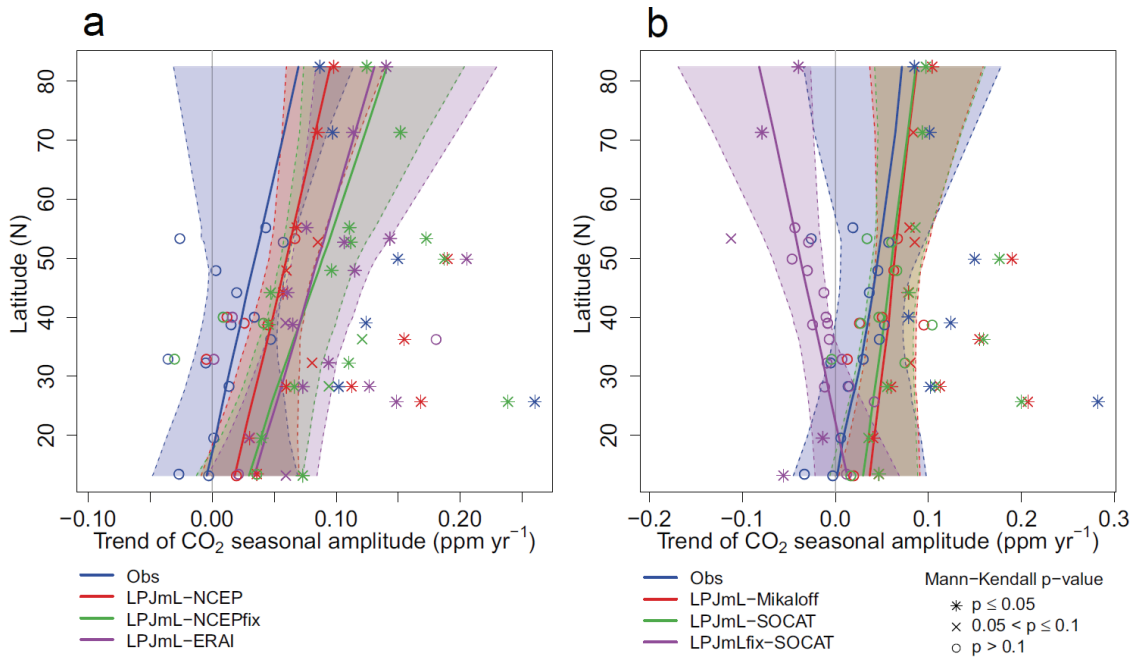


Figure 9.3: Effect of different meteorological forcing and ocean CO<sub>2</sub> uptake datasets on CO<sub>2</sub> amplitude trends.

(a) Latitudinal gradient of CO<sub>2</sub> amplitude trend depending on the meteorological forcing dataset for TM3. All simulations use NBP from the standard LPJmL model run, fossil fuel emission from EDGAR, and ocean CO<sub>2</sub> uptake from the OCLIM dataset but they differ according to the use of NCEP and ERA-Interim reanalysis datasets as drivers for TM3, respectively. Simulated CO<sub>2</sub> amplitude trends are lower with NCEP than with ERA-Interim.

The inter-annual variability of meteorological fields contributes to trends as well: CO<sub>2</sub> amplitude trends are stronger with a fixed NCEP meteorology (LPJmL-NCEPfix, fixed by repeating the data for the year 1995) than in the case of the usual NCEP meteorology including inter-annual variability. (b) Latitudinal gradient of CO<sub>2</sub> amplitude trend depending on the ocean CO<sub>2</sub> uptake dataset. The Mikaloff dataset (Mikaloff Fletcher et al., 2007) does not contain information about inter-annual variability whereas the SOCAT dataset (Rödenbeck et al., 2014) well describes the inter-annual variability of ocean CO<sub>2</sub> uptake. Nevertheless, the differences between simulated trends by using Mikaloff or SOCAT are very small. Changes in ocean CO<sub>2</sub> uptake alone cannot explain CO<sub>2</sub> amplitude trends: trends disappear if terrestrial NBP is fixed to a mean seasonal cycle and if only ocean uptake has inter-annual variability (LPJmLfix-SOCAT).

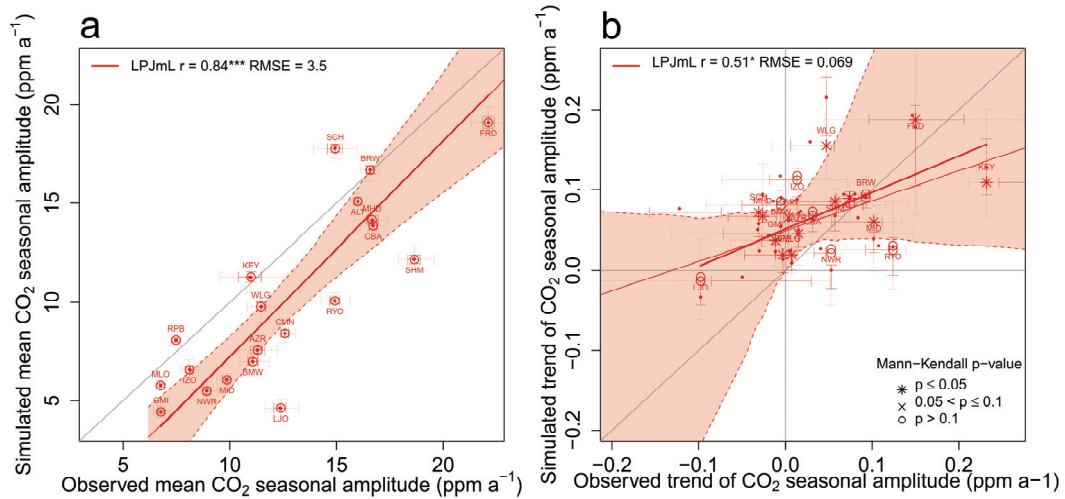


Figure 9.4: Evaluation of simulated (a) mean values and (b) trend slopes of CO<sub>2</sub> amplitude against observations.

As both x and y values have uncertainties, linear fits were computed based on orthogonal regression using principal component analysis.

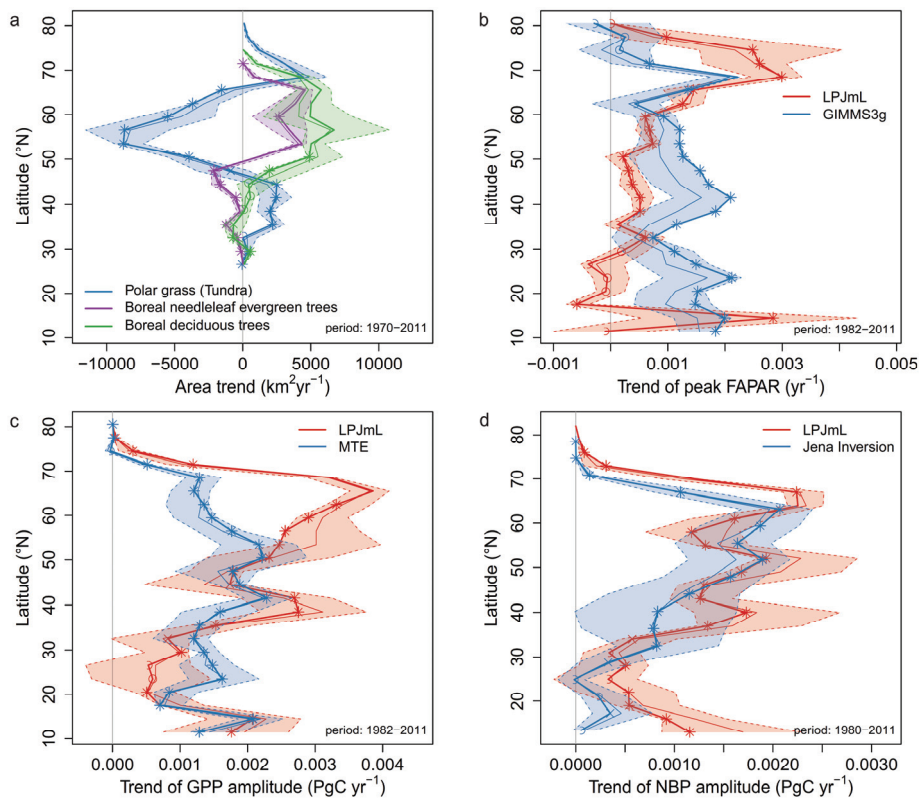


Figure 9.5: Latitudinal gradients of trends of (a) the area coverage of plant functional types as simulated with LPJmL, (b) annual maximum vegetation greenness (FAPAR, fraction of absorbed photosynthetic active radiation), (c) annual amplitude of gross primary production (GPP), and (d) of net biome productivity (NBP).

Latitudinal gradients were derived by analyzing trends on time series that aggregated over the land surface of 3° latitudinal bands. Star symbols indicate significant trends ( $p \leq 0.05$ , Mann-Kendall trend test). The Jena Inversion version s81\_v3.6 was used in (d) (using only 14 sites in inversion but covering the longest time period from 1981 onwards).



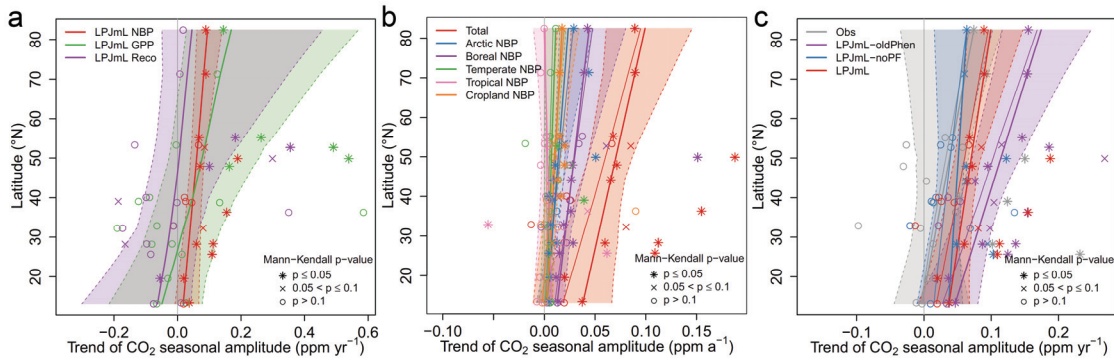


Figure 9.6: Effects of (a) terrestrial gross fluxes, (b) regions, and (c) LPJmL model components on the latitudinal gradient of CO<sub>2</sub> amplitude trends.

(a) GPP and Reco were taken from the standard LPJmL model run and transported separately within TM3. As GPP causes stronger CO<sub>2</sub> amplitude trends in northern latitudes than Reco, the latitudinal gradient of the overall CO<sub>2</sub> amplitude trend is dominated by GPP amplitude trends. (b) Boreal and arctic regions mainly shape the latitudinal gradient of CO<sub>2</sub> amplitude trends. (c) Effects of different LPJmL model setups on the latitudinal gradient of CO<sub>2</sub> amplitude trends. “LPJmL” includes permafrost dynamics and an improved phenology scheme with optimized model parameters for phenology, greenness, albedo and productivity. “LPJmL-oldPhen” uses the original LPJ phenology scheme and original model parameters and overestimate CO<sub>2</sub> amplitude trends especially in northern regions. “LPJmL-noPF” uses the new phenology scheme and optimized model parameters but the simulation of permafrost dynamics was disabled. This results in an underestimation of the slope of the latitudinal gradient.

## References

- Alcantara, C., Kuemmerle, T., Baumann, M., Bragina, E. V., Griffiths, P., Hostert, P., Knorn, J., Müller, D., Prishchepov, A. V., Schierhorn, F., Sieber, A. and Radeloff, V. C.: Mapping the extent of abandoned farmland in Central and Eastern Europe using MODIS time series satellite data, *Environ. Res. Lett.*, 8(3), 035035, doi:10.1088/1748-9326/8/3/035035, 2013.
- Alcaraz-Segura, D., Chuvieco, E., Epstein, H. E., Kasischke, E. S. and Trishchenko, A.: Debating the greening vs. browning of the North American boreal forest: differences between satellite datasets, *Glob. Change Biol.*, 16(2), 760–770, 2010.
- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H. (Ted), Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S. W., Semerci, A. and Cobb, N.: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, *For. Ecol. Manag.*, 259(4), 660–684, doi:10.1016/j.foreco.2009.09.001, 2010.
- Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Sitch, S., Piao, S. and Zhu, Z.: Evaluation of Land Surface Models in Reproducing Satellite Derived Leaf Area Index over the High-Latitude Northern Hemisphere. Part II: Earth System Models, *Remote Sens.*, 5(8), 3637–3661, doi:10.3390/rs5083637, 2013.
- Anderegg, W. R. L., Kane, J. M. and Anderegg, L. D. L.: Consequences of widespread tree mortality triggered by drought and temperature stress, *Nat. Clim. Change*, 3(1), 30–36, doi:10.1038/nclimate1635, 2013.
- Archibald, S. and Scholes, R. J.: Leaf green-up in a semi-arid African savanna – separating tree and grass responses to environmental cues, *J. Veg. Sci.*, 18(4), 583–594, 2007.
- Atzberger, C., Klisch, A., Mattiuzzi, M. and Vuolo, F.: Phenological Metrics Derived over the European Continent from NDVI3g Data and MODIS Time Series, *Remote Sens.*, 6(1), 257–284, doi:10.3390/rs6010257, 2013.
- Bacastow, R. B., Keeling, C. D. and Whorf, T. P.: Seasonal amplitude increase in atmospheric CO<sub>2</sub> concentration at Mauna Loa, Hawaii, 1959–1982, *J. Geophys. Res. Atmospheres*, 90(D6), 10529–10540, doi:10.1029/JD090iD06p10529, 1985.

- Bai, J. and Perron, P.: Computation and analysis of multiple structural change models, *J. Appl. Econom.*, 18(1), 1–22, doi:10.1002/jae.659, 2003.
- Baird, R. A. and Verbyla, D.: Browning of the landscape of interior Alaska based on 1986-2009 Landsat sensor NDVI, *Can. J. Of*, 1382, 1371–1382, doi:10.1139/x2012-088, 2012.
- Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa, S., Kozyr, A., Nojiri, Y., O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B., Wada, C., Akl, J., Barbero, L., Bates, N. R., Boutin, J., Bozec, Y., Cai, W.-J., Castle, R. D., Chavez, F. P., Chen, L., Chierici, M., Currie, K., de Baar, H. J. W., Evans, W., Feely, R. A., Fransson, A., Gao, Z., Hales, B., Hardman-Mountford, N. J., Hoppema, M., Huang, W.-J., Hunt, C. W., Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M., Jones, S. D., Jutterström, S., Kitidis, V., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Manke, A. B., Mathis, J. T., Merlivat, L., Metzl, N., Murata, A., Newberger, T., Omar, A. M., Ono, T., Park, G.-H., Paterson, K., Pierrot, D., Ríos, A. F., Sabine, C. L., Saito, S., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K. F., Sun, H., Sutton, A. J., Suzuki, T., Sweeney, C., Takahashi, T., Tjiputra, J., Tsurushima, N., van Heuven, S. M. A. C., Vandemark, D., Vlahos, P., Wallace, D. W. R., Wanninkhof, R. and Watson, A. J.: An update to the Surface Ocean CO<sub>2</sub> Atlas (SOCAT version 2), *Earth Syst. Sci. Data*, 6(1), 69–90, doi:10.5194/essd-6-69-2014, 2014.
- Baldocchi, D., Falge, E., Lianhong Gu, Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Xuhui Lee, Malhi, Y., Meyers, T., Munger, W., Oechel, W. and Paw U, K. T.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, *Bull. Am. Meteorol. Soc.*, 82(11), 2415–2434, 2001.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L. and Leroy, M.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION, *Remote Sens. Environ.*, 110(3), 275–286, doi:10.1016/j.rse.2007.02.018, 2007.
- Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P. and Smets, B.: GEOV1: LAI and FAPAR essential climate variables and FCOVER global time

- series capitalizing over existing products. Part1: Principles of development and production, *Remote Sens. Environ.*, 137, 299–309, doi:10.1016/j.rse.2012.12.027, 2013.
- Barichivich, J.: Responses of boreal vegetation to recent climate change, Doctoral thesis, University of East Anglia, May. [online] Available from: <https://ueaeprints.uea.ac.uk/49468/> (Accessed 16 January 2015), 2014.
- Barichivich, J., Briffa, K. R., Myneni, R. B., Osborn, T. J., Melvin, T. M., Ciais, P., Piao, S. and Tucker, C.: Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO<sub>2</sub> at high northern latitudes from 1950 to 2011, *Glob. Change Biol.*, 19(10), 3167–3183, doi:10.1111/gcb.12283, 2013.
- Barichivich, J., Briffa, K. R., Myneni, R., Schrier, G. van der, Dorigo, W., Tucker, C. J., Osborn, T. J. and Melvin, T. M.: Temperature and Snow-Mediated Moisture Controls of Summer Photosynthetic Activity in Northern Terrestrial Ecosystems between 1982 and 2011, *Remote Sens.*, 6(2), 1390–1431, doi:10.3390/rs6021390, 2014.
- Beck, P. S. A., Atzberger, C., Høgda, K. A., Johansen, B. and Skidmore, A. K.: Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI, *Remote Sens. Environ.*, 100(3), 321–334, doi:10.1016/j.rse.2005.10.021, 2006.
- Beck, P. S. A. and Goetz, S. J.: Satellite observations of high northern latitude vegetation productivity changes between 1982 and 2008: ecological variability and regional differences, *Environ. Res. Lett.*, 6(4), 045501, doi:10.1088/1748-9326/6/4/045501, 2011.
- Beck, P. S. A., Juday, G. P., Alix, C., Barber, V. A., Winslow, S. E., Sousa, E. E., Heiser, P., Herriges, J. D. and Goetz, S. J.: Changes in forest productivity across Alaska consistent with biome shift, *Ecol. Lett.*, 14(4), 373–9, doi:10.1111/j.1461-0248.2011.01598.x, 2011.
- Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luysaert, S., Margolis, H., Oleson, K. W., Rouspard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I. and Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, *Science*, 329(5993), 834–838, doi:10.1126/science.1184984, 2010.

- Benninghoff, W. S.: Interaction of Vegetation and Soil Frost Phenomena, *Arctic*, 5(1), 34–44, 1952.
- Berner, L. T., Beck, P. S. A., Bunn, A. G. and Goetz, S. J.: Plant response to climate change along the forest-tundra ecotone in northeastern Siberia, *Glob. Change Biol.*, 19(11), 3449–3462, doi:10.1111/gcb.12304, 2013.
- Berner, L. T., Beck, P. S. A., Bunn, A. G., Lloyd, A. H. and Goetz, S. J.: High-latitude tree growth and satellite vegetation indices: Correlations and trends in Russia and Canada (1982–2008), *J. Geophys. Res.-Biogeosciences*, 116(G01015), doi:10.1029/2010jg001475, 2011.
- De Beurs, K. M. and Henebry, G. M.: Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan, *Remote Sens. Environ.*, 89(4), 497–509, doi:10.1016/j.rse.2003.11.006, 2004a.
- De Beurs, K. M. and Henebry, G. M.: Trend Analysis of the Pathfinder AVHRR Land (PAL) NDVI Data for the Deserts of Central Asia, *IEEE Geosci. Remote Sens. Lett.*, 1(4), 282–286, 2004b.
- De Beurs, K. M. and Henebry, G. M.: A land surface phenology assessment of the northern polar regions using MODIS reflectance time series, *Can. J. REMOTE Sens.*, 36(Sp. Iss. SI Suppl. 1), S87–S110, 2010a.
- De Beurs, K. M. and Henebry, G. M.: Spatio-Temporal Statistical Methods for Modelling Land Surface Phenology, in *Phenological Research*, edited by I. L. Hudson and M. R. Keatley, pp. 177–208, Springer Netherlands. [online] Available from: [http://link.springer.com/chapter/10.1007/978-90-481-3335-2\\_9](http://link.springer.com/chapter/10.1007/978-90-481-3335-2_9) (Accessed 3 March 2014b), 2010.
- Bhatt, U. S., Walker, D. A., Raynolds, M. K., Bieniek, P. A., Epstein, H. E., Comiso, J. C., Pinzon, J. E., Tucker, C. J. and Polyakov, I. V.: Recent Declines in Warming and Vegetation Greening Trends over Pan-Arctic Tundra, *Remote Sens.*, 5(9), 4229–4254, doi:10.3390/rs5094229, 2013.
- Bi, J., Xu, L., Samanta, A., Zhu, Z. and Myneni, R.: Divergent Arctic-Boreal Vegetation Changes between North America and Eurasia over the Past 30 Years, *Remote Sens.*, 5(5), 2093–2112, doi:10.3390/rs5052093, 2013.
- Blok, D., Sass-Klaassen, U., Schaepman-Strub, G., Heijmans, M. M. P. D., Sauren, P. and Berendse, F.: What are the main climate drivers for shrub growth in

- Northeastern Siberian tundra?, *Biogeosciences*, 8(5), 1169–1179, doi:10.5194/bg-8-1169-2011, 2011a.
- Blok, D., Schaepman-Strub, G., Bartholomeus, H., Heijmans, M. M. P. D., Maximov, T. C. and Berendse, F.: The response of Arctic vegetation to the summer climate: relation between shrub cover, NDVI, surface albedo and temperature, *Environ. Res. Lett.*, 6(3), 035502, doi:10.1088/1748-9326/6/3/035502, 2011b.
- Bodegom, P. M. van, Douma, J. C. and Verheijen, L. M.: A fully traits-based approach to modeling global vegetation distribution, *Proc. Natl. Acad. Sci.*, 111(38), 13733–13738, doi:10.1073/pnas.1304551110, 2014.
- Bolstad, P. V. and Gower, S. T.: Estimation of leaf area index in fourteen southern Wisconsin forest stands using a portable radiometer, *Tree Physiol.*, 7, 115–124, 1990.
- Bonan, G.: *Ecological Climatology. Concepts and Applications.*, 2nd ed., Cambridge University Press, Cambridge., 2008a.
- Bonan, G. B.: Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests, *Science*, 320(5882), 1444–1449, doi:10.1126/science.1155121, 2008b.
- Bonan, G. B., Levis, S., Sitch, S., Vertenstein, M. and Oleson, K. W.: A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics, *Glob. Change Biol.*, 9(11), 1543–1566, doi:10.1046/j.1365-2486.2003.00681.x, 2003.
- Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M. and Smith, B.: Modelling the role of agriculture for the 20th century global terrestrial carbon balance, *Glob. Change Biol.*, (13), 679–706, 2007.
- Bradley, B. A. and Mustard, J. F.: Comparison of phenology trends by land cover class: a case study in the Great Basin, USA, *Glob. Change Biol.*, 14(2), 334–346, doi:10.1111/j.1365-2486.2007.01479.x, 2008.
- Brandt, M., Romankiewicz, C., Spiekermann, R. and Samimi, C.: Environmental change in time series – An interdisciplinary study in the Sahel of Mali and Senegal, *J. Arid Environ.*, 105, 52–63, doi:10.1016/j.jaridenv.2014.02.019, 2014a.
- Brandt, M., Verger, A., Diouf, A. A., Baret, F. and Samimi, C.: Local Vegetation Trends in the Sahel of Mali and Senegal Using Long Time Series FAPAR Satellite Products

- and Field Measurement (1982–2010), *Remote Sens.*, 6(3), 2408–2434, doi:10.3390/rs6032408, 2014b.
- Bright, B. C., Hicke, J. A. and Meddens, A. J. H.: Effects of bark beetle-caused tree mortality on biogeochemical and biogeophysical MODIS products, *J. Geophys. Res. Biogeosciences*, 118(3), 974–982, doi:10.1002/jgrg.20078, 2013.
- Broich, M., Huete, A., Tulbure, M. G., Ma, X., Xin, Q., Paget, M., Restrepo-Coupe, N., Davies, K., Devadas, R. and Held, A.: Land surface phenological response to decadal climate variability across Australia using satellite remote sensing, *Biogeosciences*, 11(18), 5181–5198, doi:10.5194/bg-11-5181-2014, 2014.
- Broyden, C. G.: The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations, *IMA J. Appl. Math.*, 6(1), 76–90, doi:10.1093/imamat/6.1.76, 1970.
- Buermann, W., Parida, B., Jung, M., MacDonald, G. M., Tucker, C. J. and Reichstein, M.: Recent Shift in Eurasian Boreal Forest Greening Response may be Associated with Warmer and Drier Summers, *Geophys. Res. Lett.*, 2014GL059450, doi:10.1002/2014GL059450, 2014.
- Buma, B., Pugh, E. T. and Wessman, C. A.: Effect of the current major insect outbreaks on decadal phenological and LAI trends in southern Rocky Mountain forests, *Int. J. Remote Sens.*, 34(20), 7249–7274, doi:10.1080/01431161.2013.817717, 2013.
- Bunn, A. G. and Goetz, S. J.: Trends in Satellite-Observed Circumpolar Photosynthetic Activity from 1982 to 2003: The Influence of Seasonality, Cover Type, and Vegetation Density, *Earth Interact.*, 10(12), 1–19, doi:10.1175/ei190.1, 2006.
- Bunn, A. G., Goetz, S. J., Kimball, J. S. and Zhang, K.: Northern High-Latitude Ecosystems Respond to Climate Change, *Eos Trans. Am. Geophys. Union*, 88(34), 333–333, doi:10.1029/2007eo340001, 2007.
- Burnham, K. P. and Anderson, D. R.: Model selection and multimodel inference a practical information-theoretic approach, Springer, New York. [online] Available from: <http://site.ebrary.com/id/10047705> (Accessed 17 December 2013), 2002.
- Von Buttler, J., Zscheischler, J. and Mahecha, M. D.: An extended approach for spatiotemporal gapfilling: dealing with large and systematic gaps in geoscientific datasets, *Nonlin Process. Geophys.*, 21(1), 203–215, doi:10.5194/npg-21-203-2014, 2014.

- Caffarra, A. and Donnelly, A.: The ecological significance of phenology in four different tree species: effects of light and temperature on bud burst, *Int. J. Biometeorol.*, 55(5), 711–721, doi:10.1007/s00484-010-0386-1, 2011.
- Cai, W., Yuan, W., Liang, S., Liu, S., Dong, W., Chen, Y., Liu, D. and Zhang, H.: Large Differences in Terrestrial Vegetation Production Derived from Satellite-Based Light Use Efficiency Models, *Remote Sens.*, 6(9), 8945–8965, doi:10.3390/rs6098945, 2014.
- Caldararu, S., Palmer, P. I. and Purves, D. W.: Inferring Amazon leaf demography from satellite observations of leaf area index, *Biogeosciences*, 9(4), 1389–1404, doi:10.5194/bg-9-1389-2012, 2012.
- Caldararu, S., Purves, D. W. and Palmer, P. I.: Phenology as a strategy for carbon optimality: a global model, *Biogeosciences*, 11(3), 763–778, doi:10.5194/bg-11-763-2014, 2014.
- Carlson, T. N. and Ripley, D. A.: On the relation between NDVI, fractional vegetation cover, and leaf area index, *Remote Sens. Environ.*, 62(3), 241–252, doi:10.1016/S0034-4257(97)00104-1, 1997.
- Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T. and Reichstein, M.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, *Nature*, 514(7521), 213–217, doi:10.1038/nature13731, 2014.
- Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S. and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, *Glob. Change Biol.*, 16(10), 2813–2829, doi:10.1111/j.1365-2486.2010.02173.x, 2010.
- Cescatti, A., Marcolla, B., Santhana Vannan, S. K., Pan, J. Y., Román, M. O., Yang, X., Ciais, P., Cook, R. B., Law, B. E., Matteucci, G., Migliavacca, M., Moors, E., Richardson, A. D., Seufert, G. and Schaaf, C. B.: Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network, *Remote Sens. Environ.*, 121, 323–334, doi:10.1016/j.rse.2012.02.019, 2012.



- CFS: Canadian Large Fire Database. Natural Resources Canada, Canadian Forest Service, edited by C. F. Service, [online] Available from: [http://cwfis.cfs.nrcan.gc.ca/en\\_CA/nfdbID - 427](http://cwfis.cfs.nrcan.gc.ca/en_CA/nfdbID - 427), 2010.
- Chapin, F. S., Sturm, M., Serreze, M. C., McFadden, J. P., Key, J. R., Lloyd, A. H., McGuire, A. D., Rupp, T. S., Lynch, A. H., Schimel, J. P., Beringer, J., Chapman, W. L., Epstein, H. E., Euskirchen, E. S., Hinzman, L. D., Jia, G., Ping, C.-L., Tape, K. D., Thompson, C. D. C., Walker, D. A. and Welker, J. M.: Role of Land-Surface Changes in Arctic Summer Warming, *Science*, 310(5748), 657–660, doi:10.1126/science.1117368, 2005.
- Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., Wirth, C., Aber, J. D., Cole, J. J., Goulden, M. L., Harden, J. W., Heimann, M., Howarth, R. W., Matson, P. A., McGuire, A. D., Melillo, J. M., Mooney, H. A., Neff, J. C., Houghton, R. A., Pace, M. L., Ryan, M. G., Running, S. W., Sala, O. E., Schlesinger, W. H. and Schulze, E.-D.: Reconciling Carbon-cycle Concepts, Terminology, and Methods, *Ecosystems*, 9(7), 1041–1050, doi:10.1007/s10021-005-0105-7, 2006.
- Chen, J. M., Menges, C. H. and Leblanc, S. G.: Global mapping of foliage clumping index using multi-angular satellite data, *Remote Sens. Environ.*, 97(4), 447–457, doi:10.1016/j.rse.2005.05.003, 2005.
- Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V. and Ladurie, E. L. R.: Historical phenology: Grape ripening as a past climate indicator, *Nature*, 432(7015), 289–290, doi:10.1038/432289a, 2004.
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J. G., Chhabra, A., DeFries, R. S., Galloway, J., Heimann, Martin, Jones, C., Le Quéré, C., Myneni, R., Piao, S. and Thornton, P. E.: Carbon and Other Biogeochemical Cycles, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, J. Xia, Bex, V., and Midgley, P. M., pp. 465–570, Cambridge, United Kingdom and New York, USA. [online] Available from: <http://www.climatechange2013.org/report/full-report/>, 2013.

- Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A. and Schwartz, M. D.: Shifting plant phenology in response to global change, *Trends Ecol. Evol.*, 22(7), 357–365, doi:10.1016/j.tree.2007.04.003, 2007.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E. and Terpenning, I.: STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *J. Off. Stat.*, 6(1), 3–73, 1990.
- Congalton, R. G.: A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data, *Remote Sens. Environ.*, 37, 35–46, 1991.
- Congalton, R. G., Gu, J., Yadav, K., Thenkabail, P. and Ozdogan, M.: Global Land Cover Mapping: A Review and Uncertainty Analysis, *Remote Sens.*, 6(12), 12070–12093, doi:10.3390/rs61212070, 2014.
- Conway, T. J., Tans, P. P., Waterman, L. S., Thoning, K. W., Kitzis, D. R., Masarie, K. A. and Zhang, N.: Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res. Atmospheres*, 99(D11), 22831–22855, doi:10.1029/94JD01951, 1994.
- Cook, B. I. and Pau, S.: A Global Assessment of Long-Term Greening and Browning Trends in Pasture Lands Using the GIMMS LAI3g Dataset, *Remote Sens.*, 5(5), 2492–2512, doi:10.3390/rs5052492, 2013.
- Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. and Lambin, E.: Digital change detection methods in ecosystem monitoring: a review, *Int. J. Remote Sens.*, 25(9), 1565–1596, doi:10.1080/0143116031000101675, 2004.
- Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., Kucharik, C., Lomas, M. R., Ramankutty, N., Sitch, S., Smith, B., White, A. and Young-Molling, C.: Global response of terrestrial ecosystem structure and function to CO<sub>2</sub> and climate change: results from six dynamic global vegetation models, *Glob. Change Biol.*, 7(4), 357–373, doi:10.1046/j.1365-2486.2001.00383.x, 2001.
- Dardel, C., Kergoat, L., Hiernaux, P., Grippa, M., Mougou, E., Ciais, P. and Nguyen, C.-C.: Rain-Use-Efficiency: What it Tells us about the Conflicting Sahel Greening and Sahelian Paradox, *Remote Sens.*, 6(4), 3446–3474, doi:10.3390/rs6043446, 2014a.

- Dardel, C., Kergoat, L., Hiernaux, P., Mougin, E., Grippa, M. and Tucker, C. J.: Re-greening Sahel: 30 years of remote sensing data and field observations (Mali, Niger), *Remote Sens. Environ.*, 140, 350–364, doi:10.1016/j.rse.2013.09.011, 2014b.
- Davison, J. E., Breshears, D. D., van Leeuwen, W. J. D. and Casady, G. M.: Remotely sensed vegetation phenology and productivity along a climatic gradient: on the value of incorporating the dimension of woody plant cover, *Glob. Ecol. Biogeogr.*, 20(1), 101–113, doi:10.1111/j.1466-8238.2010.00571.x, 2011.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J. N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137(656), 553–597, doi:10.1002/qj.828, 2011.
- DeFries, R., Hansen, M. and Townshend, J.: Global discrimination of land cover types from metrics derived from AVHRR pathfinder data, *Remote Sens. Environ.*, 54(3), 209–222, doi:10.1016/0034-4257(95)00142-5, 1995.
- Delbart, N., Le Toan, T., Kergoat, L. and Fedotova, V.: Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004), *Remote Sens. Environ.*, 101(1), 52–62, doi:10.1016/j.rse.2005.11.012, 2006.
- Deming, W. E. and Stephan, F. F.: On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known, *Ann. Math. Stat.*, 11(4), 427–444, doi:10.1214/aoms/1177731829, 1940.
- D’Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E. and Schaepman, M. E.: Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe, *Remote Sens. Environ.*, 142, 141–154, doi:10.1016/j.rse.2013.12.005, 2014.
- Donohue, R. J., Roderick, M. L., McVicar, T. R. and Farquhar, G. D.: Impact of CO<sub>2</sub> fertilization on maximum foliage cover across the globe’s warm, arid environments, *Geophys. Res. Lett.*, 40(12), 3031–3035, doi:10.1002/grl.50563, 2013.

- Dorigo, W. A., Zurita-Milla, R., Wit, A. J. W. de, Brazile, J., Singh, R. and Schaepman, M. E.: A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling, *Int. J. Appl. Earth Obs. Geoinformation*, 9(2), 165–193, 2007.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, *Remote Sens. Environ.*, 120, 25–36, doi:10.1016/j.rse.2011.11.026, 2012.
- Edburg, S. L., Hicke, J. A., Brooks, P. D., Pendall, E. G., Ewers, B. E., Norton, U., Gochis, D., Gutmann, E. D. and Meddens, A. J.: Cascading impacts of bark beetle-caused tree mortality on coupled biogeophysical and biogeochemical processes, *Front. Ecol. Environ.*, 10(8), 416–424, doi:10.1890/110173, 2012.
- Edburg, S. L., Hicke, J. A., Lawrence, D. M. and Thornton, P. E.: Simulating coupled carbon and nitrogen dynamics following mountain pine beetle outbreaks in the western United States, *J. Geophys. Res. Biogeosciences*, 116(G4), G04033, doi:10.1029/2011JG001786, 2011.
- Eklundh, L., Johansson, T. and Solberg, S.: Mapping insect defoliation in Scots pine with MODIS time-series data, *Remote Sens. Environ.*, 113(7), 1566–1573, doi:10.1016/j.rse.2009.03.008, 2009.
- Eklundh, L. and Olsson, L.: Vegetation index trends for the African Sahel 1982–1999, *Geophys. Res. Lett.*, 30(8), 1–4, doi:10.1029/2002gl016772, 2003.
- Elmore, A. J., Guinn, S. M., Minsley, B. J. and Richardson, A. D.: Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests, *Glob. Change Biol.*, 18(2), 656–674, doi:10.1111/j.1365-2486.2011.02521.x, 2012.
- Emerson, R. and Lewis, C. M.: The Dependence of the Quantum Yield of Chlorella Photosynthesis on Wave Length of Light, *Am. J. Bot.*, 30(3), 165–178, doi:10.2307/2437236, 1943.
- Engelmann, T. W.: Die Erscheinungsweise der Sauerstoffausscheidung chromophyllhaltiger Zellen im Licht bei Anwendung der Bacterienmethode, *Arch. Für Gesamte Physiol. Menschen Tiere*, 57(8-9), 375–386, doi:10.1007/BF01662192, 1894.

- Fader, M., Rost, S., Müller, C., Bondeau, A. and Gerten, D.: Virtual water content of temperate cereals and maize: Present and potential future patterns, *J. Hydrol.*, 384(3–4), 218–231, doi:10.1016/j.jhydrol.2009.12.011, 2010.
- Fensholt, R. and Proud, S. R.: Evaluation of Earth Observation based global long term vegetation trends – Comparing GIMMS and MODIS global NDVI time series, *Remote Sens. Environ.*, 119, 131–147, doi:10.1016/j.rse.2011.12.015, 2012.
- Fensholt, R., Rasmussen, K., Kaspersen, P., Huber, S., Horion, S. and Swinnen, E.: Assessing Land Degradation/Recovery in the African Sahel from Long-Term Earth Observation Based Primary Productivity and Precipitation Relationships, *Remote Sens.*, 5(2), 664–686, doi:10.3390/rs5020664, 2013.
- Fensholt, R., Rasmussen, K., Nielsen, T. T. and Mbow, C.: Evaluation of earth observation based long term vegetation trends – Intercomparing NDVI time series trend analysis consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT data, *Remote Sens. Environ.*, 113(9), 1886–1898, doi:10.1016/j.rse.2009.04.004, 2009.
- Fensholt, R., Sandholt, I. and Rasmussen, M. S.: Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements, *Remote Sens. Environ.*, 91(3-4), 490–507, doi:10.1016/j.rse.2004.04.009, 2004.
- Fisher, J. B., Badgley, G. and Blyth, E.: Global nutrient limitation in terrestrial vegetation, *Glob. Biogeochem. Cycles*, 26(3), GB3007, doi:10.1029/2011GB004252, 2012.
- Fletcher, R.: A new approach to variable metric algorithms, *Comput. J.*, 13(3), 317–322, doi:10.1093/comjnl/13.3.317, 1970.
- Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N. and Snyder, P. K.: Global Consequences of Land Use, *Science*, 309(5734), 570–574, doi:10.1126/science.1111772, 2005.
- Foley, J. A., Levis, S., Costa, M. H., Cramer, W. and Pollard, D.: Incorporating Dynamic Vegetation Cover within Global Climate Models, *Ecol. Appl.*, 10(6), 1620–1632, doi:10.2307/2641227, 2000.

- Foody, G. M.: Status of land cover classification accuracy assessment, *Remote Sens. Environ.*, 80(1), 185–201, doi:10.1016/s0034-4257(01)00295-4, 2002.
- Forbes, B. C., Macias Fauria, M. and Zetterberg, P.: Russian Arctic warming and 'greening' are closely tracked by tundra shrub willows, *Glob. Change Biol.*, 16(5), 1542–1554, 2010.
- Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.
- Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M., Neigh, C. and Reichstein, M.: Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology, *Remote Sens.*, 5(5), 2113–2144, doi:10.3390/rs5052113, 2013.
- Forkel, M., Thonicke, K., Beer, C., Cramer, W., Bartalev, S. and Schmullius, C.: Extreme fire events are related to previous-year surface moisture conditions in permafrost-underlain larch forests of Siberia, *Environ. Res. Lett.*, 7(4), 044021, doi:10.1088/1748-9326/7/4/044021, 2012.
- Forsythe, G. E., Malcolm, M. A. and Moler, C. B.: Computer methods for mathematical computations, Prentice-Hall., 1977.
- Frames: Alaska Large Fire Database, [online] Available from: <http://www.frames.gov/rcs/10000/10465.html>, 2012.
- Francey, R. J., Steele, L.P., Spencer, D.A., Langenfelds, R.L., Law, R.M., Krummel, P.B., Fraser, P.J., Etheridge, D.M., Derek, N., Coram, S.A., Cooper, L.N., Allison, C.E., Porter, L. and Baly, S.: The CSIRO (Australia) measurement of greenhouse gases in the global atmosphere, in Report of the 11th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, pp. 97–111, World Meteorological Organization Global Atmosphere Watch., 2003.
- Fraser, R. H., Lantz, T. C., Olthof, I., Kokelj, S. V. and Sims, R. A.: Warming-Induced Shrub Expansion and Lichen Decline in the Western Canadian Arctic, *Ecosystems*, 17(7), 1151–1168, doi:10.1007/s10021-014-9783-3, 2014.
- Fraser, R. H., Olthof, I., Carrière, M., Deschamps, a and Pouliot, D.: Detecting long-term changes to vegetation in northern Canada using the Landsat satellite image

- archive, *Environ. Res. Lett.*, 6(4), 045502–045502, doi:10.1088/1748-9326/6/4/045502, 2011.
- Fridley, J. D.: Extended leaf phenology and the autumn niche in deciduous forest invasions, *Nature*, 485(7398), 359–362, doi:10.1038/nature11056, 2012.
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C. and Zeng, N.: Climate–Carbon Cycle Feedback Analysis: Results from the C<sup>4</sup> MIP Model Intercomparison, *J. Clim.*, 19(14), 3337–3353, doi:10.1175/JCLI3800.1, 2006.
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K. and Knutti, R.: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, *J. Clim.*, 27(2), 511–526, doi:10.1175/JCLI-D-12-00579.1, 2014.
- Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A. and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO<sub>2</sub>, *Proc. Natl. Acad. Sci.*, 111(9), 3280–3285, doi:10.1073/pnas.1222477110, 2014.
- Frost, G. V., Epstein, H. E. and Walker, D. A.: Regional and landscape-scale variability of Landsat-observed vegetation dynamics in northwest Siberian tundra, *Environ. Res. Lett.*, 9(2), 025004, doi:10.1088/1748-9326/9/2/025004, 2014.
- Fuchs, R., Herold, M., Verburg, P. H. and Clevers, J. G. P. W.: A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe, *Biogeosciences*, 10(3), 1543–1559, doi:10.5194/bg-10-1543-2013, 2013.
- Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. P. W. and Eberle, J.: Gross changes in reconstructions of historic land cover/use for Europe between 1900–2010, *Glob. Change Biol.*, n/a–n/a, doi:10.1111/gcb.12714, 2014.
- Fu, Y. H., Piao, S., Op de Beeck, M., Cong, N., Zhao, H., Zhang, Y., Menzel, A. and Janssens, I. A.: Recent spring phenology shifts in western Central Europe based on

- multiscale observations, *Glob. Ecol. Biogeogr.*, 23(11), 1255–1263, doi:10.1111/geb.12210, 2014a.
- Fu, Y. H., Piao, S., Zhao, H., Jeong, S.-J., Wang, X., Vitisse, Y., Ciais, P. and Janssens, I. A.: Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern middle and high latitudes, *Glob. Change Biol.*, 20(12), 3743–3755, doi:10.1111/gcb.12610, 2014b.
- Fu, Y. S. H., Campioli, M., Vitisse, Y., Boeck, H. J. D., Berge, J. V. den, AbdElgawad, H., Asard, H., Piao, S., Deckmyn, G. and Janssens, I. A.: Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species, *Proc. Natl. Acad. Sci.*, 111(20), 7355–7360, doi:10.1073/pnas.1321727111, 2014c.
- Galvez, D. A., Landhäusser, S. m. and Tyree, M. t.: Root carbon reserve dynamics in aspen seedlings: does simulated drought induce reserve limitation?, *Tree Physiol.*, 31(3), 250–257, 2011.
- Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, E., Joel, G., Peñuelas, J. and Valentini, R.: Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types, *Ecol. Appl.*, 5(1), 28–41, 1995.
- Garonna, I., de Jong, R., de Wit, A. J. W., Mùcher, C. A., Schmid, B. and Schaepman, M. E.: Strong contribution of autumn phenology to changes in satellite-derived growing season length estimates across Europe (1982–2011), *Glob. Change Biol.*, 20(11), 3457–3470, doi:10.1111/gcb.12625, 2014.
- Geng, L., Ma, M., Wang, X., Yu, W., Jia, S. and Wang, H.: Comparison of Eight Techniques for Reconstructing Multi-Satellite Sensor Time-Series NDVI Data Sets in the Heihe River Basin, China, *Remote Sens.*, 6(3), 2024–2049, doi:10.3390/rs6032024, 2014.
- Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W. and Sitch, S.: Terrestrial vegetation and water balance - hydrological evaluation of a dynamic global vegetation model, *J. Hydrol.*, (286), 249–270, 2004.
- Gervois, S., de Noblet-Ducoudré, N., Viovy, N., Ciais, P., Brisson, N., Seguin, B. and Perrier, A.: Including Croplands in a Global Biosphere Model: Methodology and Evaluation at Specific Sites, *Earth Interact.*, 8(16), 1–25, doi:10.1175/1087-3562(2004)8<1:ICIAGB>2.0.CO;2, 2004.



- Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C. and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, *Biogeosciences*, 7(3), 1171–1186, 2010.
- Gobron, N., Pinty, B., Aussenat, O., Chen, J. M., Cohen, W. B., Fensholt, R., Gond, V., Huemmrich, K. F., Lavergne, T., Mélin, F., Privette, J. L., Sandholt, I., Taberner, M., Turner, D. P., Verstraete, M. M. and Widlowski, J.-L.: Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations, *J. Geophys. Res.*, 111(D13), D13110–D13110, doi:10.1029/2005jd006511, 2006.
- Gobron, N., Pinty, B., Aussenat, O., Taberner, M., Faber, O., Mélin, F., Lavergne, T., Robustelli, M. and Snoeij, P.: Uncertainty estimates for the FAPAR operational products derived from MERIS — Impact of top-of-atmosphere radiance uncertainties and validation with field data, *Remote Sens. Environ.*, 112(4), 1871–1883, doi:10.1016/j.rse.2007.09.011, 2008.
- Gobron, N. and Verstraete, M. M.: FAPAR Fraction of Absorbed Photosynthetically Active Radiation, Global Terrestrial Observing System, Rome. [online] Available from: <http://www.fao.org/gtos/doc/ECVs/T10/T10.pdf>, 2009.
- Goetz, S. J., Bunn, A. G., Fiske, G. J. and Houghton, R. A.: Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance, *Proc. Natl. Acad. Sci. U. S. A.*, 102(38), 13521–13525, doi:10.1073/pnas.0506179102, 2005.
- Goldfarb, D.: A family of variable-metric methods derived by variational means, *Math. Comput.*, 24(109), 23–26, doi:10.1090/S0025-5718-1970-0258249-6, 1970.
- Golyandina, N., Nekrutkin, V. V. and Zhigljavskiy, A. A.: Analysis of time series structure: SSA and related techniques, Chapman & Hall/CRC, Boca Raton, Fla. [online] Available from: <http://www.loc.gov/catdir/enhancements/fy0646/00050442-d.html>, 2001.
- Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., Welp, L. R., Sweeney, C., Tans, P. P., Kelley, J. J., Daube, B. C., Kort, E. A., Santoni, G. W. and Bent, J. D.: Enhanced Seasonal Exchange of CO<sub>2</sub> by Northern Ecosystems Since 1960, *Science*, 341(6150), 1085–1089, doi:10.1126/science.1239207, 2013.

- Gray, J. M., Frolking, S., Kort, E. A., Ray, D. K., Kucharik, C. J., Ramankutty, N. and Friedl, M. A.: Direct human influence on atmospheric CO<sub>2</sub> seasonality from increased cropland productivity, *Nature*, 515(7527), 398–401, doi:10.1038/nature13957, 2014.
- Guay, K. C., Beck, P. S. A., Berner, L. T., Goetz, S. J., Baccini, A. and Buermann, W.: Vegetation productivity patterns at high northern latitudes: a multi-sensor satellite data assessment, *Glob. Change Biol.*, 20(10), 3147–3158, doi:10.1111/gcb.12647, 2014.
- Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.
- Hamunyela, E., Verbesselt, J., Roerink, G. and Herold, M.: Trends in Spring Phenology of Western European Deciduous Forests, *Remote Sens.*, 5(12), 6159–6179, doi:10.3390/rs5126159, 2013.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, *Science*, 342(6160), 850–853, doi:10.1126/science.1244693, 2013.
- Harris, I., Jones, P. d., Osborn, T. j. and Lister, D. h.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, *Int. J. Climatol.*, 34(3), 623–642, doi:10.1002/joc.3711, 2014.
- Hartmann, H., Adams, H. D., Anderegg, W. R. L., Jansen, S. and Zeppel, M. J. B.: Research frontiers in drought-induced tree mortality: crossing scales and disciplines, *New Phytol.*, 205(3), 965–969, doi:10.1111/nph.13246, 2015.
- Hastie, T. J. and Tibshirani, R. J.: *Generalized Additive Models*, CRC Press., 1990.
- Heimann, M., Keeling, C. D. and Tucker, C. J.: A three dimensional model of atmospheric CO<sub>2</sub> transport based on observed winds: 3. Seasonal cycle and synoptic time scale variations, in *Aspects of Climate Variability in the Pacific and the Western Americas*, edited by D. H. Peterson, pp. 277–303, American Geophysical Union. [online] Available from:

- <http://onlinelibrary.wiley.com/doi/10.1029/GM055p0277/summary> (Accessed 17 October 2014), 1989.
- Herrmann, S. M., Anyamba, A. and Tucker, C. J.: Recent trends in vegetation dynamics in the African Sahel and their relationship to climate, *Glob. Environ. Change*, 15(4), 394–404, doi:10.1016/j.gloenvcha.2005.08.004, 2005.
- Hess, J. C., Scott, C. A., Hufford, G. L. and Fleming, M. D.: El Nino and its impact on fire weather conditions in Alaska, *Int. J. Wildland Fire*, 10(1), 1–13, doi:10.1071/wf01007, 2001.
- Hickler, T., Eklundh, L., Seaquist, J. W., Smith, B., Ardö, J., Olsson, L., Sykes, M. T. and Sjöström, M.: Precipitation controls Sahel greening trend, *Geophys. Res. Lett.*, 32(21), L21415, doi:10.1029/2005GL024370, 2005.
- Hilker, T., Natsagdorj, E., Waring, R. H., Lyapustin, A. and Wang, Y.: Satellite observed widespread decline in Mongolian grasslands largely due to overgrazing, *Glob. Change Biol.*, 20(2), 418–428, doi:10.1111/gcb.12365, 2014.
- Hird, J. N. and McDermid, G. J.: Noise reduction of NDVI time series: An empirical comparison of selected techniques, *Remote Sens. Environ.*, 113(1), 248–258, doi:10.1016/j.rse.2008.09.003, 2009.
- Hlásny, T., Zajíčková, L., Turčáni, M., Holuša, J. and Sitková, Z.: Geographical variability of spruce bark beetle development under climate change in the Czech Republic, *J. For. Sci.*, 57, 242–249, 2011.
- Høgda, K. A., Karlsen, S. R. and Solheim, I.: Climatic change impact on growing season in Fennoscandia studied by a time series of NOAA AVHRR NDVI data, in *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*, vol. 3, pp. 1338–1340 vol.3., 2001.
- Høgda, K. A., Tømmervik, H. and Karlsen, S. R.: Trends in the Start of the Growing Season in Fennoscandia 1982–2011, *Remote Sens.*, 5(9), 4304–4318, doi:10.3390/rs5094304, 2013.
- Holben, B. N.: Characteristics of maximum-value composite images from temporal AVHRR data, *Int. J. Remote Sens.*, 7(11), 1417–1434, 1986.
- Hollinger, D. Y., Ollinger, S. V., Richardson, a D., Meyers, T. P., Dail, D. B., Martin, M. E., Scott, N. a, Arkebauer, T. J., Baldocchi, D. D., Clark, K. L., Curtis, P. S., Davis, K. J., Desai, a R., Dragoni, D., Goulden, M. L., Gu, L., Katul, G. G., Pallardy, S. G., Paw

- U, K. T., Schmid, H. P., Stoy, P. C., Suyker, a E. and Verma, S. B.: Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration, *Glob. Change Biol.*, 16(2), 696–710, doi:10.1111/j.1365-2486.2009.02028.x, 2010.
- Horion, S., Fensholt, R., Tagesson, T. and Ehammer, A.: Using earth observation-based dry season NDVI trends for assessment of changes in tree cover in the Sahel, *Int. J. Remote Sens.*, 35(7), 2493–2515, doi:10.1080/01431161.2014.883104, 2014.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ.*, 83(1–2), 195–213, doi:10.1016/S0034-4257(02)00096-2, 2002.
- Huete, A. R.: Vegetation Indices, Remote Sensing and Forest Monitoring, *Geogr. Compass*, 6(9), 513–532, doi:10.1111/j.1749-8198.2012.00507.x, 2012.
- Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutyyra, L. R., Yang, W., Nemani, R. R. and Myneni, R.: Amazon rainforests green-up with sunlight in dry season, *Geophys. Res. Lett.*, 33(6), L06405, doi:10.1029/2005GL025583, 2006.
- Huete, A. R., Hua, G., Qi, J., Chehbouni, A. and van Leeuwen, W. J. D.: Normalization of multidirectional red and NIR reflectances with the SAVI, *Remote Sens. Environ.*, 41(2–3), 143–154, doi:10.1016/0034-4257(92)90074-T, 1992.
- Humboldt, A. von: *Kosmos: Entwurf einer physischen Weltbeschreibung*, Eichborn, Frankfurt am Main., 1845.
- Huntington, T. G.: Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, 319(1–4), 83–95, doi:10.1016/j.jhydrol.2005.07.003, 2006.
- Hyvönen, R., Ågren, G. I., Linder, S., Persson, T., Cotrufo, M. F., Ekblad, A., Freeman, M., Grelle, A., Janssens, I. A., Jarvis, P. G., Kellomäki, S., Lindroth, A., Loustau, D., Lundmark, T., Norby, R. J., Oren, R., Pilegaard, K., Ryan, M. G., Sigurdsson, B. D., Strömgren, M., van Oijen, M. and Wallin, G.: The likely impact of elevated [CO<sub>2</sub>], nitrogen deposition, increased temperature and management on carbon sequestration in temperate and boreal forest ecosystems: a literature review, *New Phytol.*, 173(3), 463–480, doi:10.1111/j.1469-8137.2007.01967.x, 2007.

- IPCC: Climate Change 2013 - The Physical Science Basis - Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge., 2014.
- Ito, A.: A historical meta-analysis of global terrestrial net primary productivity: are estimates converging?, *Glob. Change Biol.*, 17(10), 3161–3175, doi:10.1111/j.1365-2486.2011.02450.x, 2011.
- Janssen, P. H. M. and Heuberger, P. S. C.: Calibration of process-oriented models, *Ecol. Model.*, (83), 55–66, 1995.
- Jarvis, P. and Linder, S.: Botany: Constraints to growth of boreal forests, *Nature*, 405(6789), 904–905, doi:10.1038/35016154, 2000.
- Jeong, J.-H., Kug, J.-S., Linderholm, H. W., Chen, D., Kim, B.-M. and Jun, S.-Y.: Intensified Arctic warming under greenhouse warming by vegetation–atmosphere–sea ice interaction, *Environ. Res. Lett.*, 9(9), 094007, doi:10.1088/1748-9326/9/9/094007, 2014.
- Jeong, S.-J., Ho, C.-H., Gim, H.-J. and Brown, M. E.: Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008, *Glob. Change Biol.*, no–no, doi:10.1111/j.1365-2486.2011.02397.x, 2011.
- Jia, G. J., Epstein, H. E. and Walker, D. A.: Vegetation greening in the canadian arctic related to decadal warming, *J. Environ. Monit.*, 11(12), 2231–2238, 2009.
- Jiang, N., Zhu, W., Zheng, Z., Chen, G. and Fan, D.: A Comparative Analysis between GIMSS NDVIg and NDVI3g for Monitoring Vegetation Activity Change in the Northern Hemisphere during 1982–2008, *Remote Sens.*, 5(8), 4031–4044, doi:10.3390/rs5084031, 2013.
- Jiang, Y. and Huang, B.: Physiological Responses to Heat Stress Alone or in Combination with Drought: A Comparison between Tall Fescue and Perennial Ryegrass, *HortScience*, 36(4), 682–686, 2001.
- Jolly, W. M., Nemani, R. and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, *Glob. Change Biol.*, 11(4), 619–632, doi:10.1111/j.1365-2486.2005.00930.x, 2005.

- De Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E. and Dent, D. L.: Analysis of monotonic greening and browning trends from global NDVI time-series, *Remote Sens. Environ.*, 115(2), 692–702, doi:10.1016/j.rse.2010.10.011, 2011a.
- De Jong, R., Schaepman, M. E., Furrer, R., de Bruin, S. and Verburg, P. H.: Spatial relationship between climatologies and changes in global vegetation activity, *Glob. Change Biol.*, 19(6), 1953–1964, doi:10.1111/gcb.12193, 2013a.
- De Jong, R., Verbesselt, J., Schaepman, M. E. and Bruin, S.: Trend changes in global greening and browning: contribution of short-term trends to longer-term change, *Glob. Change Biol.*, 18(2), 642–655, doi:10.1111/j.1365-2486.2011.02578.x, 2011b.
- De Jong, R., Verbesselt, J., Zeileis, A. and Schaepman, M.: Shifts in Global Vegetation Activity Trends, *Remote Sens.*, 5(3), 1117–1133, doi:10.3390/rs5031117, 2013b.
- Jönsson, A. M., Appelberg, G., Harding, S. and Barring, L.: Spatio-temporal impact of climate change on the activity and voltinism of the spruce bark beetle, *Ips typographus*, *Glob. Change Biol.*, 15(2), 486–499, doi:10.1111/j.1365-2486.2008.01742.x, 2009.
- Jonsson, P. and Eklundh, L.: Seasonality extraction by function fitting to time-series of satellite sensor data, *IEEE Trans. Geosci. Remote Sens.*, 40(8), 1824–1832, doi:10.1109/TGRS.2002.802519, 2002.
- Jorgenson, M. T., Romanovsky, V. E., Harden, J. W., Shur, Y. L., O'Donnell, J., Schuur, E. A. G., Kanevskiy, M. and Marchenko, S.: Resilience and vulnerability of permafrost to climate change, *Can. J. For. Res. - Rev. Can. Rech. For.*, 40(7), 1219–1236, 2010.
- Julien, Y. and Sobrino, J. A.: Global land surface phenology trends from GIMMS database, *Int. J. Remote Sens.*, 30(13), 3495–3513, doi:10.1080/01431160802562255, 2009.
- Jung, M., Henkel, K., Herold, M. and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sens. Environ.*, 101, 534–553, doi:10.1016/j.rse.2006.01.020, 2006.
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F. and Williams, C.: Global

- patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.*, 116, G00J07–G00J07, doi:10.1029/2010jg001566, 2011.
- Jung, M., Verstraete, M., Gobron, N., Reichstein, M., Papale, D., Bondeau, A., Robustelli, M. and Pinty, B.: Diagnostic assessment of European gross primary production, *Glob. Change Biol.*, 14(10), 2349–2364, doi:10.1111/j.1365-2486.2008.01647.x, 2008.
- Justice, C. O., Townshend, J., Holben, B. N. and Tucker, C.: Analysis of the phenology of global vegetation using meteorological satellite data, *Int. J. Remote Sens.*, 6(8), 1271–1318, doi:10.1080/01431168508948281, 1985.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, 77(3), 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRPP>2.0.CO;2, 1996.
- Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R. and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO<sub>2</sub> into a terrestrial vegetation model and interactive mission benefit analysis, *Biogeosciences*, 9(8), 3173–3184, doi:10.5194/bg-9-3173-2012, 2012.
- Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H. and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, *J. Geophys. Res. Biogeosciences*, 118(4), 1414–1426, doi:10.1002/jgrg.20118, 2013.
- Kandasamy, S., Baret, F., Verger, A., Neveux, P. and Weiss, M.: A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products, *Biogeosciences*, 10(6), 4055–4071, doi:10.5194/bg-10-4055-2013, 2013.

- Karlsen, S. R., Elvebakk, A., Høgda, K. A. and Johansen, B.: Satellite-based mapping of the growing season and bioclimatic zones in Fennoscandia, *Glob. Ecol. Biogeogr.*, 15(4), 416–430, doi:10.1111/j.1466-822X.2006.00234.x, 2006.
- Kasischke, E. S., Verbyla, D. L., Rupp, T. S., McGuire, A. D., Murphy, K. A., Jandt, R., Barnes, J. L., Hoy, E. E., Duffy, P. A., Calef, M. and Turetsky, M. R.: Alaska's changing fire regime - implications for the vulnerability of its boreal forests, *Can. J. For. Res. - Rev. Can. Rech. For.*, 40(7), 1313–1324, 2010.
- Kasischke, E. S., Williams, D. and Barry, D.: Analysis of the patterns of large fires in the boreal forest region of Alaska, *Int. J. Wildland Fire*, 11(2), 131–144, 2002.
- Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., Atkin, O., Bahn, M., Baker, T. R., Baldocchi, D., Bekker, R., Blanco, C. C., Blonder, B., Bond, W. J., Bradstock, R., Bunker, D. E., Casanoves, F., Cavender-Bares, J., Chambers, J. Q., Chapin Iii, F. S., Chave, J., Coomes, D., Cornwell, W. K., Craine, J. M., Dobrin, B. H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W. F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G. T., Fyllas, N. M., Gallagher, R. V., Green, W. A., Gutierrez, A. G., Hickler, T., Higgins, S. I., Hodgson, J. G., Jalili, A., Jansen, S., Joly, C. A., Kerkhoff, A. J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J. M. H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T. D., Leishman, M., Lens, F., Lenz, T., Lewis, S. L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M. D., Manning, P., Massad, T., Medlyn, B. E., Messier, J., Moles, A. T., Müller, S. C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V. G., Onoda, Y., Ordoñez, J., Overbeck, G., et al.: TRY – a global database of plant traits, *Glob. Change Biol.*, 17(9), 2905–2935, doi:10.1111/j.1365-2486.2011.02451.x, 2011.
- Kaufmann, R. K., Zhou, L., Knyazikhin, Y., Shabanov, V., Myneni, R. B. and Tucker, C. J.: Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data, *IEEE Trans. Geosci. Remote Sens.*, 38(6), 2584–2597, doi:10.1109/36.885205, 2000.



- Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R., Waterman, L. S. and Chin, J. F. S.: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus*, 28(6), 538–551, doi:10.1111/j.2153-3490.1976.tb00701.x, 1976.
- Keeling, C. D., Chin, J. F. S. and Whorf, T. P.: Increased activity of northern vegetation inferred from atmospheric CO<sub>2</sub> measurements, *Nature*, 382(6587), 146–149, doi:10.1038/382146a0, 1996.
- Keeling, R. F., Piper, S. C., Bollenbacher, A. F. and Walker, S. J.: Atmospheric CO<sub>2</sub> records from sites in the SIO air sampling network. Trends: A Compendium of Data on Global Change, [online] Available from: <http://cdiac.ornl.gov/trends/co2/sio-keel.html>, 2009.
- Keenan, T., Carbone, M., Reichstein, M. and Richardson, A.: The model–data fusion pitfall: assuming certainty in an uncertain world, *Oecologia*, 167(3), 587–597, doi:10.1007/s00442-011-2106-x, 2011.
- Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O’Keefe, J., Schmid, H. P., Wing, I. S., Yang, B. and Richardson, A. D.: Net carbon uptake has increased through warming-induced changes in temperate forest phenology, *Nat. Clim. Change*, 4(7), 598–604, doi:10.1038/nclimate2253, 2014.
- Kehlmann, D.: *Die Vermessung der Welt*, 38th ed., Rowolth Verlag GmbH, Reinbek., 2007.
- Kelley, D. I., Prentice, I. C., Harrison, S. P., Wang, H., Simard, M., Fisher, J. B. and Willis, K. O.: A comprehensive benchmarking system for evaluating global vegetation models, *Biogeosciences*, 10(5), 3313–3340, doi:10.5194/bg-10-3313-2013, 2013.
- Kendall, M. G.: *Rank Correlation Methods*, Griffin, London., 1975.
- Keuper, F., van Bodegom, P. M., Dorrepaal, E., Weedon, J. T., van Hal, J., van Logtestijn, R. S. P. and Aerts, R.: A frozen feast: thawing permafrost increases plant-available nitrogen in subarctic peatlands, *Glob. Change Biol.*, 18(6), 1998–2007, doi:10.1111/j.1365-2486.2012.02663.x, 2012.
- Kira, T., Shinokazi, K. and Hozumi, K.: Structure of forest canopies as related to their primary productivity, *Plant Cell Physiol.*, 10(1), 129–142, 1969.

- Kleidon, A., Fraedrich, K. and Heimann, M.: A Green Planet Versus a Desert World: Estimating the Maximum Effect of Vegetation on the Land Surface Climate, *Clim. Change*, 44(4), 471–493, doi:10.1023/A:1005559518889, 2000.
- Klosterman, S. T., Hufkens, K., Gray, J. M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M. A. and Richardson, A. D.: Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery, *Biogeosciences*, 11(16), 4305–4320, doi:10.5194/bg-11-4305-2014, 2014.
- Kloster, S., Mahowald, N. M., Randerson, J. T., Thornton, P. E., Hoffman, F. M., Levis, S., Lawrence, P. J., Feddes, J. J., Oleson, K. W. and Lawrence, D. M.: Fire dynamics during the 20th century simulated by the Community Land Model, *Biogeosciences*, 7(6), 1877–1902, doi:10.5194/bg-7-1877-2010, 2010.
- Knorr, W.: Annual and interannual CO<sub>2</sub> exchanges of the terrestrial biosphere: process-based simulations and uncertainties, *Glob. Ecol. Biogeogr.*, 9(3), 225–252, doi:10.1046/j.1365-2699.2000.00159.x, 2000.
- Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R. and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, *J. Geophys. Res. Biogeosciences*, 115(G4), G04017, doi:10.1029/2009JG001119, 2010.
- Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, P., Myneni, R. B., Nemani, R. and Running, S. W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document. [online] Available from: <http://eosps0.gsfc.nasa.gov/atbd/modistables.html>, 1999.
- Koffi, E., Rayner, P., Scholze, M., Chevallier, F. and Kaminski, T.: Quantifying the constraint of biospheric process parameters by CO<sub>2</sub> concentration and flux measurement networks through a carbon cycle data assimilation system, *Atmos Chem Phys Discuss*, 12(9), 24131–24172, doi:10.5194/acpd-12-24131-2012, 2012.
- Köppen, W.: Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit und nach der Wirkung der Wärme auf die organische Welt betrachtet, *Meteorol. Z.*, 1, 215–226, 1884.
- Köppen, W. and Geiger, R.: *Klima der Erde*, 1954.

- Körner, C. and Basler, D.: Phenology Under Global Warming, *Science*, 327(5972), 1461–1462, doi:10.1126/science.1186473, 2010.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, *Meteorol. Z.*, 15(3), 259–263, doi:10.1127/0941-2948/2006/0130, 2006.
- Kozak, J.: Forest Cover Change in the Western Carpathians in the Past 180 Years, *Mt. Res. Dev.*, 23(4), 369–375, doi:10.1659/0276-4741(2003)023[0369:FCCITW]2.0.CO;2, 2003.
- Kramer, K., Leinonen, I. and Loustau, D.: The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and Mediterranean forests ecosystems: an overview, *Int. J. Biometeorol.*, 44(2), 67–75, 2000.
- Krause, P., Boyle, D. P. and Bäse, F.: Comparison of different efficiency criteria for hydrological model assessment, *Adv Geosci*, 5, 89–97, doi:10.5194/adgeo-5-89-2005, 2005.
- Kubíková, J.: Forest dieback in Czechoslovakia, *Vegetatio*, 93(2), 101–108, doi:10.1007/BF00033204, 1991.
- Kucharik, C. J.: Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield, *Earth Interact.*, 7(14), 1–33, doi:10.1175/1087-3562(2003)007<0001:EOAPAM>2.0.CO;2, 2003.
- Kuemmerle, T., Hostert, P., Radeloff, V. C., Linden, S. van der, Perzanowski, K. and Kruhlov, I.: Cross-border Comparison of Post-socialist Farmland Abandonment in the Carpathians, *Ecosystems*, 11(4), 614–628, doi:10.1007/s10021-008-9146-z, 2008.
- Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F. and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*, 9(10), 3757–3776, doi:10.5194/bg-9-3757-2012, 2012.
- Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L. and Cescatti, A.: Model–data fusion across ecosystems: from multisite optimizations to global simulations, *Geosci Model Dev*, 7(6), 2581–2597, doi:10.5194/gmd-7-2581-2014, 2014.

- Kurz, W. A., Dymond, C. C., Stinson, G., Rampley, G. J., Neilson, E. T., Carroll, A. L., Ebata, T. and Safranyik, L.: Mountain pine beetle and forest carbon feedback to climate change, *Nature*, 452(7190), 987–990, 2008.
- Ladysz, J.: Chosen aspects of sustainable development on the Polish, Czech and German borderland, *GeoJournal*, 67(1), 1–7, doi:10.1007/s10708-006-9002-7, 2006.
- Larcher, W.: *Physiological Plant Ecology. Ecophysiology and Stress Physiology of Functional Groups*, 4th ed., Springer, Berlin., 2003.
- LeBauer, D. S. and Treseder, K. K.: Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed, *Ecology*, 89(2), 371–379, doi:10.1890/06-2057.1, 2008.
- Van Leeuwen, W., Hartfield, K., Miranda, M. and Meza, F.: Trends and ENSO/AAO Driven Variability in NDVI Derived Productivity and Phenology alongside the Andes Mountains, *Remote Sens.*, 5(3), 1177–1203, doi:10.3390/rs5031177, 2013.
- Lehsten, V., Tansey, K. J., Balzter, H., Thonicke, K., Spessa, A., Weber, U., Smith, B. and Arneeth, A.: Estimating carbon emissions from African wildfires, *Biogeosciences Discuss*, 5(4), 3091–3122, doi:10.5194/bgd-5-3091-2008, 2008.
- Levenberg, K.: A method for the solution of certain non-linear problems in least squares, *Q. J. Appl. Mathematics*, II(2), 164–168, 1944.
- Levin, I., Graul, R. and Trivett, N. B. A.: Long-term observations of atmospheric CO<sub>2</sub> and carbon isotopes at continental sites in Germany, *Tellus B*, 47(1-2), 23–34, doi:10.1034/j.1600-0889.47.issue1.4.x, 1995.
- Liang, L., Schwartz, M. D. and Fei, S.: Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest, *Remote Sens. Environ.*, 115(1), 143–157, doi:10.1016/j.rse.2010.08.013, 2011.
- Lieth, H.: Purposes of a Phenology Book, in *Phenology and Seasonality Modeling*, edited by H. Lieth, pp. 3–19, Springer Berlin Heidelberg. [online] Available from: [http://link.springer.com/chapter/10.1007/978-3-642-51863-8\\_1](http://link.springer.com/chapter/10.1007/978-3-642-51863-8_1) (Accessed 25 September 2014), 1974.
- Lieth, H. H.: Contributions to phenology seasonality research, *Int. J. Biometeorol.*, 20(3), 197–199, doi:10.1007/BF01553661, 1976.

- Liu, H., Tian, F., Hu, H. C., Hu, H. P. and Sivapalan, M.: Soil moisture controls on patterns of grass green-up in Inner Mongolia : an index based approach, *Hydrol. Earth Syst. Sci.*, 17(2004), 805–815, doi:10.5194/hess-17-805-2013, 2013.
- Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, 43(7), W07401, doi:10.1029/2006WR005756, 2007.
- Loranty, M. M., Berner, L. T., Goetz, S. J., Jin, Y. and Randerson, J. T.: Vegetation controls on northern high latitude snow-albedo feedback: observations and CMIP5 model simulations, *Glob. Change Biol.*, 20(2), 594–606, doi:10.1111/gcb.12391, 2014.
- Loranty, M. M., Goetz, S. J. and Beck, P. S. A.: Tundra vegetation effects on pan-Arctic albedo, *Environ. Res. Lett.*, 6(2), 024014, doi:10.1088/1748-9326/6/2/024014, 2011.
- Lucht, W., Prentice, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W. and Smith, B.: Climatic Control of the High-Latitude Vegetation Greening Trend and Pinatubo Effect, *Science*, 296(5573), 1687–1689, doi:10.1126/science.1071828, 2002.
- Lucht, W., Schaaf, C. B. and Strahler, A. H.: An algorithm for the retrieval of albedo from space using semiempirical BRDF models, *Geosci. Remote Sens. IEEE Trans. On*, 38(2), 977–998, doi:10.1109/36.841980, 2000.
- Lucht, W., Schaphoff, S., Erbrecht, T., Heyder, U. and Cramer, W.: Terrestrial vegetation redistribution and carbon balance under climate change, *Carbon Balance Manag.*, 1, 6–6, doi:10.1186/1750-0680-1-6, 2006.
- Macias Fauria, M. and Johnson, E. a: Large-scale climatic patterns control large lightning fire occurrence in Canada and Alaska forest regions, *J. Geophys. Res.*, 111(G4), doi:10.1029/2006jg000181, 2006.
- Madani, N., Kimball, J. S., Affleck, D. L. R., Kattge, J., Graham, J., van Bodegom, P. M., Reich, P. B. and Running, S. W.: Improving ecosystem productivity modeling through spatially explicit estimation of optimal light use efficiency, *J. Geophys. Res. Biogeosciences*, 119(9), 2014JG002709, doi:10.1002/2014JG002709, 2014.
- Maeno, H. and Hiura, T.: The effect of leaf phenology of overstory trees on the reproductive success of an understory shrub, *Staphylea bumalda* DC, *Can. J. Bot.*, 78(6), 781–785, doi:10.1139/b00-053, 2000.

- Mahecha, M. D., Fürst, L. M., Gobron, N. and Lange, H.: Identifying multiple spatiotemporal patterns: A refined view on terrestrial photosynthetic activity, *Pattern Recognit. Lett.*, 31(14), 2309–2317, 2010a.
- Mahecha, M. D., Reichstein, M., Jung, M., Seneviratne, S. I., Zaehle, S., Beer, C., Braakhekke, M. C., Carvalhais, N., Lange, H., Le Maire, G. and Moors, E.: Comparing observations and process-based simulations of biosphere-atmosphere exchanges on multiple timescales, *J. Geophys. Res.*, 115(G2), G02003–G02003, doi:10.1029/2009jg001016, 2010b.
- Maignan, F., Bréon, F. M., Chevallier, F., Viovy, N., Ciais, P., Garrec, C., Trules, J. and Mancip, M.: Evaluation of a Global Vegetation Model using time series of satellite vegetation indices, *Geosci Model Dev*, 4(4), 1103–1114, doi:10.5194/gmd-4-1103-2011, 2011.
- Main-Knorn, M., Hostert, P., Kozak, J. and Kuemmerle, T.: How pollution legacies and land use histories shape post-communist forest cover trends in the Western Carpathians, *For. Ecol. Manag.*, 258(2), 60–70, doi:10.1016/j.foreco.2009.03.034, 2009.
- Mann, H. B.: Nonparametric Tests Against Trend, *Econometrica*, 13(3), 245–259, 1945.
- Mao, J., Shi, X., Thornton, P. E., Hoffman, F. M., Zhu, Z. and Myneni, R. B.: Global Latitudinal-Asymmetric Vegetation Growth Trends and Their Driving Mechanisms: 1982–2009, *Remote Sens.*, 5(3), 1484–1497, doi:10.3390/rs5031484, 2013.
- Mao, J., Shi, X., Thornton, P. E., Piao, S. and Wang, X.: Causes of spring vegetation growth trends in the northern mid–high latitudes from 1982 to 2004, *Environ. Res. Lett.*, 7(1), 014010–014010, doi:10.1088/1748-9326/7/1/014010, 2012.
- Markert, B., Herpin, U., Berlekamp, J., Oehlmann, J., Grodzinska, K., Mankovska, B., Suchara, I., Siewers, U., Weckert, V. and Lieth, H.: A comparison of heavy metal deposition in selected Eastern European countries using the moss monitoring method, with special emphasis on the “Black Triangle,” *Sci. Total Environ.*, 193(2), 85–100, doi:10.1016/S0048-9697(96)05333-8, 1996.
- Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *J. Soc. Ind. Appl. Math.*, 11(2), 431–441, 1963.
- Mathieu, P.-P. and O’Neill, A.: Data assimilation: From photon counts to Earth System forecasts: Remote Sensing Data Assimilation Special Issue, *Remote Sens. Environ.*, 112(4), 1258–1267, 2008.

- Maurer, C., Koch, E., Hammerl, C., Hammerl, T. and Pokorny, E.: BACCHUS temperature reconstruction for the period 16th to 18th centuries from Viennese and Klosterneuburg grape harvest dates, *J. Geophys. Res. Atmospheres*, 114(D22), D22106, doi:10.1029/2009JD011730, 2009.
- McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S. and Nilsson, S.: Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000, *Remote Sens. Environ.*, 114(5), 941–949, doi:10.1016/j.rse.2009.12.009, 2010.
- McDowell, N. G., Beerling, D. J., Breshears, D. D., Fisher, R. A., Raffa, K. F. and Stitt, M.: The interdependence of mechanisms underlying climate-driven vegetation mortality, *Trends Ecol. Evol.*, 26(10), 523–532, doi:10.1016/j.tree.2011.06.003, 2011.
- McKinley, G. A., Fay, A. R., Takahashi, T. and Metzl, N.: Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales, *Nat. Geosci.*, 4(9), 606–610, doi:10.1038/ngeo1193, 2011.
- McMichael, C. E., Hope, A. S., Stow, D. A. and Fleming, J. B.: The relation between active layer depth and a spectral vegetation index in arctic tundra landscapes of the North Slope of Alaska, *Int. J. Remote Sens.*, 18(11), 2371–2382, doi:10.1080/014311697217666, 1997.
- McPhaden, M. J.: Genesis and Evolution of the 1997-98 El Niño, *Science*, 283(5404), 950–954, doi:10.1126/science.283.5404.950, 1999.
- Mebane, W. R. and Sekhon, J. S.: Genetic Optimization Using Derivatives: The rgenoud Package for R, *J. Stat. Softw.*, 42(11), 2011.
- Melillo, J. M., McGuire, A. D., Kicklighter, D. W., Moore, B., Vorosmarty, C. J. and Schloss, A. L.: Global climate change and terrestrial net primary production, *Nature*, 363(6426), 234–240, doi:10.1038/363234a0, 1993.
- Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen, P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A. J. H., Wielgolaski, F.-E., Zach, S. and Züst, A.: European phenological response to climate change matches the warming pattern, *Glob. Change Biol.*, 12(10), 1969–1976, doi:10.1111/j.1365-2486.2006.01193.x, 2006.

- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E.: Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, 21(6), 1087–1092, doi:10.1063/1.1699114, 1953.
- Miao, L., Luan, Y., Luo, X., Liu, Q., Moore, J. C., Nath, R., He, B., Zhu, F. and Cui, X.: Analysis of the Phenology in the Mongolian Plateau by Inter-Comparison of Global Vegetation Datasets, *Remote Sens.*, 5(10), 5193–5208, doi:10.3390/rs5105193, 2013.
- Migliavacca, M., Cremonese, E., Colombo, R., Busetto, L., Galvagno, M., Ganis, L., Meroni, M., Pari, E., Rossini, M., Siniscalco, C. and Morra di Cella, U.: European larch phenology in the Alps: can we grasp the role of ecological factors by combining field observations and inverse modelling?, *Int. J. Biometeorol.*, 52(7), 587–605, doi:10.1007/s00484-008-0152-9, 2008.
- Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., Cogliati, S., Manca, G., Diotri, F., Busetto, L., Cescatti, A., Colombo, R., Fava, F., Morra di Cella, U., Pari, E., Siniscalco, C. and Richardson, A. D.: Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO<sub>2</sub> uptake, *Agric. For. Meteorol.*, 151(10), 1325–1337, doi:10.1016/j.agrformet.2011.05.012, 2011.
- Migliavacca, M., Meroni, M., Busetto, L., Colombo, R., Zenone, T., Matteucci, G., Manca, G. and Seufert, G.: Modeling Gross Primary Production of Agro-Forestry Ecosystems by Assimilation of Satellite-Derived Information in a Process-Based Model, *Sensors*, 9(2), 922–942, doi:10.3390/s90200922, 2009.
- Migliavacca, M., Sonnentag, O., Keenan, T. F., Cescatti, A., O’Keefe, J. and Richardson, A. D.: On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model, *Biogeosciences*, 9(6), 2063–2083, doi:10.5194/bg-9-2063-2012, 2012.
- Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Gloor, M., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A. and Sarmiento, J. L.: Inverse estimates of the oceanic sources and sinks of natural CO<sub>2</sub> and the implied oceanic carbon transport, *Glob. Biogeochem. Cycles*, 21(1), GB1010, doi:10.1029/2006GB002751, 2007.
- Mitchell, M.: *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, USA, 1998.



- Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, 25(6), 693–712, doi:10.1002/joc.1181, 2005.
- Molau, U.: Responses to natural climatic variation and experimental warming in two tundra plant species with contrasting life forms: *Cassiope tetragona* and *Ranunculus nivalis*, *Glob. Change Biol.*, 3(S1), 97–107, doi:10.1111/j.1365-2486.1997.gcb138.x, 1997.
- Monsi, M. and Saeki, T.: Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion, *Jpn. J. Bot.*, 14, 22–52, 1953.
- Monsi, M. and Saeki, T.: On the Factor Light in Plant Communities and its Importance for Matter Production, *Ann. Bot.*, 95(3), 549–567, doi:10.1093/aob/mci052, 2005.
- Monteith, J. L.: Solar Radiation and Productivity in Tropical Ecosystems, *J. Appl. Ecol.*, 9(3), 747, doi:10.2307/2401901, 1972.
- Morain, S. A.: Phenology and Remote Sensing, in *Phenology and Seasonality Modeling*, edited by H. Lieth, pp. 55–75, Springer Berlin Heidelberg. [online] Available from: [http://link.springer.com/chapter/10.1007/978-3-642-51863-8\\_5](http://link.springer.com/chapter/10.1007/978-3-642-51863-8_5) (Accessed 25 September 2014), 1974.
- Morton, D. C., Nagol, J., Carabajal, C. C., Rosette, J., Palace, M., Cook, B. D., Vermote, E. F., Harding, D. J. and North, P. R. J.: Amazon forests maintain consistent canopy structure and greenness during the dry season, *Nature*, 506, 221–224, doi:10.1038/nature13006, 2014.
- Mudelsee, M.: Break function regression, *Eur. Phys. J. Spec. Top.*, 174(1), 49–63, doi:10.1140/epjst/e2009-01089-3, 2009.
- Mueller, T., Dressler, G., Tucker, C. J., Pinzon, J. E., Leimgruber, P., Dubayah, R. O., Hurtt, G. C., Böhning-Gaese, K. and Fagan, W. F.: Human Land-Use Practices Lead to Global Long-Term Increases in Photosynthetic Capacity, *Remote Sens.*, 6(6), 5717–5731, doi:10.3390/rs6065717, 2014.
- Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Ahlström, A., Lomas, M., Levis, S., Viovy, N. and Zeng, N.: Evaluation of Land Surface Models in Reproducing Satellite-Derived LAI over the High-Latitude Northern Hemisphere. Part I: Uncoupled DGVMs, *Remote Sens.*, 5(10), 4819–4838, doi:10.3390/rs5104819, 2013.

- Musial, J. P., Verstraete, M. M. and Gobron, N.: Technical Note: Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series, *Atmos Chem Phys*, 11(15), 7905–7923, doi:10.5194/acp-11-7905-2011, 2011.
- Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. a, Schmidt, N. M., Schaepman-Strub, G., Wipf, S., Rixen, C., Ménard, C. B., Venn, S., Goetz, S., Andreu-Hayles, L., Elmendorf, S., Ravolainen, V., Welker, J., Grogan, P., Epstein, H. E. and Hik, D. S.: Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities, *Environ. Res. Lett.*, 6(4), 045509–045509, doi:10.1088/1748-9326/6/4/045509, 2011.
- Myneni, R. B., Hall, F. G., Sellers, P. J. and Marshak, A. L.: The interpretation of spectral vegetation indexes, *Geosci. Remote Sens. IEEE Trans. On*, 33(2), 481–486, doi:10.1109/36.377948, 1995.
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. and Nemani, R. R.: Increased plant growth in the northern high latitudes from 1981 to 1991, *Nature*, 386(6626), 698–702, doi:10.1038/386698a0, 1997a.
- Myneni, R. B., Ramakrishna, R., Nemani, R. and Running, S. W.: Estimation of global leaf area index and absorbed par using radiative transfer models, *IEEE Trans. Geosci. Remote Sens.*, 35(6), 1380–1393, doi:10.1109/36.649788, 1997b.
- Myneni, R. B. and Williams, D. L.: On the relationship between FAPAR and NDVI, *Remote Sens. Environ.*, 49(3), 200–211, doi:10.1016/0034-4257(94)90016-7, 1994.
- Myneni, R. B., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., Didan, K., Fu, R., Juárez, R. I. N., Saatchi, S. S., Hashimoto, H., Ichii, K., Shabanov, N. V., Tan, B., Ratana, P., Privette, J. L., Morisette, J. T., Vermote, E. F., Roy, D. P., Wolfe, R. E., Friedl, M. A., Running, S. W., Votava, P., El-Saleous, N., Devadiga, S., Su, Y. and Salomonson, V. V.: Large seasonal swings in leaf area of Amazon rainforests, *Proc. Natl. Acad. Sci.*, 104(12), 4820–4823, doi:10.1073/pnas.0611338104, 2007.
- NASA: Mission Details “Landsat Science,” *Landsat Sci.* [online] Available from: [http://landsat.gsfc.nasa.gov/?page\\_id=4091](http://landsat.gsfc.nasa.gov/?page_id=4091) (Accessed 10 January 2015), 2015.

- Natali, S. M., Schuur, E. A. G. and Rubin, R. L.: Increased plant productivity in Alaskan tundra as a result of experimental warming of soil and permafrost, *J. Ecol.*, 100(2), 488–498, doi:10.1111/j.1365-2745.2011.01925.x, 2012.
- Neigh, C. S. R., Bolton, D. K., Diabate, M., Williams, J. J. and Carvalhais, N.: An Automated Approach to Map the History of Forest Disturbance from Insect Mortality and Harvest with Landsat Time-Series Data, *Remote Sens.*, 6(4), 2782–2808, doi:10.3390/rs6042782, 2014.
- Neigh, C., Tucker, C. and Townshend, J.: North American vegetation dynamics observed with multi-resolution satellite data, *Remote Sens. Environ.*, 112(4), 1749–1772, doi:10.1016/j.rse.2007.08.018, 2008.
- Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., Myneni, R. B. and Running, S. W.: Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999, *Science*, 300(5625), 1560–1563, doi:10.1126/science.1082750, 2003.
- Nepstad, D. C., Tohver, I. M., Ray, D., Moutinho, P. and Cardinot, G.: Mortality of Large Trees and Lianas Following Experimental Drought in an Amazon Forest, *Ecology*, 88(9), 2259–2269, doi:10.1890/06-1046.1, 2007.
- New, M., Todd, M., Hulme, M. and Jones, P.: Precipitation measurements and trends in the twentieth century, *Int. J. Climatol.*, 21(15), 1889–1922, doi:10.1002/joc.680, 2001.
- Niinemets, Ü., Keenan, T. F. and Hallik, L.: A worldwide analysis of within-canopy variations in leaf structural, chemical and physiological traits across plant functional types, *New Phytol.*, 205(3), 973–993, doi:10.1111/nph.13096, 2015.
- Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., Ledford, J., McCarthy, H. R., Moore, D. J. P., Ceulemans, R., Angelis, P. D., Finzi, A. C., Karnosky, D. F., Kubiske, M. E., Lukac, M., Pregitzer, K. S., Scarascia-Mugnozza, G. E., Schlesinger, W. H. and Oren, R.: Forest response to elevated CO<sub>2</sub> is conserved across a broad range of productivity, *Proc. Natl. Acad. Sci. U. S. A.*, 102(50), 18052–18056, doi:10.1073/pnas.0509478102, 2005.
- Ohta, T., Maximov, T. C., Dolman, A. J., Nakai, T., van der Molen, M. K., Kononov, A. V., Maximov, A. P., Hiyama, T., Iijima, Y., Moors, E. J., Tanaka, H., Toba, T. and Yabuki, H.: Interannual variation of water balance and summer evapotranspiration

- in an eastern Siberian larch forest over a 7-year period (1998-2006), *Agric. For. Meteorol.*, 148(12), 1941–1953, 2008.
- Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in *The Climate system*, pp. 33–78., 2001.
- Olsson, L., Eklundh, L. and Ardö, J.: A recent greening of the Sahel—trends, patterns and potential causes, *J. Arid Environ.*, 63(3), 556–566, doi:10.1016/j.jaridenv.2005.03.008, 2005.
- Olsson, P.-O., Jönsson, A. M. and Eklundh, L.: A new invasive insect in Sweden – *Physokermes inopinatus*: Tracing forest damage with satellite based remote sensing, *For. Ecol. Manag.*, 285, 29–37, doi:10.1016/j.foreco.2012.08.003, 2012.
- Papale, D. and Valentini, R.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, *Glob. Change Biol.*, 9(4), 525–535, doi:10.1046/j.1365-2486.2003.00609.x, 2003.
- Parent, M. B. and Verbyla, D.: The Browning of Alaska’s Boreal Forest, *Remote Sens.*, 2(12), 2729–2747, 2010.
- Pavlick, R., Drewry, D. T., Bohn, K., Reu, B. and Kleidon, A.: The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs, *Biogeosciences*, 10(6), 4137–4177, doi:10.5194/bg-10-4137-2013, 2013.
- Van Peer, L., Nijs, I., Reheul, D. and De Cauwer, B.: Species richness and susceptibility to heat and drought extremes in synthesized grassland ecosystems: compositional vs physiological effects, *Funct. Ecol.*, 18(6), 769–778, doi:10.1111/j.0269-8463.2004.00901.x, 2004.
- Peñuelas, J., Rutishauser, T. and Filella, I.: Phenology Feedbacks on Climate Change, *Science*, 324(5929), 887–888, doi:10.1126/science.1173004, 2009.
- Phillips, O. L., Aragão, L. E. O. C., Lewis, S. L., Fisher, J. B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C. A., Heijden, G. van der, Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T. R., Bánki, O., Blanc, L., Bonal, D., Brando, P., Chave, J., Oliveira, Á. C. A. de, Cardozo, N. D., Czimczik, C. I., Feldpausch, T. R., Freitas, M. A., Gloor, E., Higuchi, N., Jiménez, E., Lloyd, G., Meir, P., Mendoza, C., Morel, A., Neill, D. A., Nepstad, D., Patiño, S., Peñuela, M.

- C., Prieto, A., Ramírez, F., Schwarz, M., Silva, J., Silveira, M., Thomas, A. S., Steege, H. ter, Stropp, J., Vásquez, R., Zelazowski, P., Dávila, E. A., Andelman, S., Andrade, A., Chao, K.-J., Erwin, T., Fiore, A. D., C. E. H., Keeling, H., Killeen, T. J., Laurance, W. F., Cruz, A. P., Pitman, N. C. A., Vargas, P. N., Ramírez-Angulo, H., Rudas, A., Salamão, R., Silva, N., Terborgh, J. and Torres-Lezama, A.: Drought Sensitivity of the Amazon Rainforest, *Science*, 323(5919), 1344–1347, doi:10.1126/science.1164033, 2009.
- Piao, S., Friedlingstein, P., Ciais, P., Zhou, L. and Chen, A.: Effect of climate and CO<sub>2</sub> changes on the greening of the Northern Hemisphere over the past two decades, *Geophys. Res. Lett.*, 33(23), L23402, doi:10.1029/2006GL028205, 2006.
- Piao, S., Nan, H., Huntingford, C., Ciais, P., Friedlingstein, P., Sitch, S., Peng, S., Ahlström, A., Canadell, J. G., Cong, N., Levis, S., Levy, P. E., Liu, L., Lomas, M. R., Mao, J., Myneni, R. B., Peylin, P., Poulter, B., Shi, X., Yin, G., Viovy, N., Wang, T., Wang, X., Zaehle, S., Zeng, N., Zeng, Z. and Chen, A.: Evidence for a weakening relationship between interannual temperature variability and northern vegetation activity, *Nat. Commun.*, 5, doi:10.1038/ncomms6018, 2014.
- Piao, S., Wang, X., Ciais, P., Zhu, B., Wang, T. and Liu, J.: Changes in satellite-derived vegetation growth trend in temperate and boreal Eurasia from 1982 to 2006, *Glob. Change Biol.*, 17(10), 3228–3239, doi:10.1111/j.1365-2486.2011.02419.x, 2011.
- Pinty, B., Lavergne, T., Voßbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M. M. and Widlowski, J.-L.: Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products, *J. Geophys. Res. Atmospheres*, 112(D10), D10116, doi:10.1029/2006JD008105, 2007.
- Pinzon, J. E. and Tucker, C. J.: A Non-Stationary 1981–2012 AVHRR NDVI3g Time Series, *Remote Sens.*, 6(8), 6929–6960, doi:10.3390/rs6086929, 2014.
- Plummer, S., Arino, O., Ranera, F., Tansey, K., Chen, J., Dedieu, G., Eva, H., Piccolini, I., Leigh, R., Borstlap, G., Beusen, B., Heyns, W. and Benedetti, R.: The GLOBCARBON Initiative: Global Biophysical Products for Terrestrial Carbon Studies, *IEEE Int. Geosci. Remote Sens. Symp. IGARSS 2007 Barc.*, 2007, 2408–2411, 2007.

- Poirier, M., Durand, J.-L. and Volaire, F.: Persistence and production of perennial grasses under water deficits and extreme temperatures: importance of intraspecific vs. interspecific variability, *Glob. Change Biol.*, 18(12), 3632–3646, doi:10.1111/j.1365-2486.2012.02800.x, 2012.
- Polák, T., Albrechtová, J. and Rock, B. N.: Bud development types as a new macroscopic marker of Norway spruce decline and recovery processes along a mountainous pollution gradient, *Forestry*, 79(4), 425–437, doi:10.1093/forestry/cpl009, 2006.
- Porada, P., Weber, B., Elbert, W., Pöschl, U. and Kleidon, A.: Estimating global carbon uptake by lichens and bryophytes with a process-based model, *Biogeosciences*, 10(11), 6989–7033, doi:10.5194/bg-10-6989-2013, 2013.
- Potter, C. S., Klooster, S. and Brooks, V.: Interannual Variability in Terrestrial Net Primary Production: Exploration of Trends and Controls on Regional to Global Scales, *Ecosystems*, 2(1), 36–48, doi:10.1007/s100219900056, 1999.
- Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S. and Zimmermann, N. E.: Plant functional type mapping for earth system models, *Geosci Model Dev*, 4(4), 993–1010, doi:10.5194/gmd-4-993-2011, 2011a.
- Poulter, B., Frank, D. C., Hodson, E. L. and Zimmermann, N. E.: Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO<sub>2</sub> airborne fraction, *Biogeosciences*, 8(8), 2027–2036, doi:10.5194/bg-8-2027-2011, 2011b.
- Prentice, I. C., Bondeau, A., Cramer, W., Harrison, S. P., Hickler, T., Lucht, W., Sitch, S., Smith, B. and Sykes, M. T.: Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change, in *Terrestrial Ecosystems in a Changing World*, edited by J. G. Canadell, D. E. Pataki, and L. F. Pitelka, pp. 175–192, Springer Berlin Heidelberg. [online] Available from: [http://link.springer.com/chapter/10.1007/978-3-540-32730-1\\_15](http://link.springer.com/chapter/10.1007/978-3-540-32730-1_15) (Accessed 6 January 2014), 2007.
- Prentice, I. C., Kelley, D. I., Foster, P. N., Friedlingstein, P., Harrison, S. P. and Bartlein, P. J.: Modeling fire and the terrestrial carbon balance, *Glob. Biogeochem. Cycles*, 25(3), GB3005, doi:10.1029/2010GB003906, 2011.

- Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B. E., Disney, M. and Bowyer, P.: Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter, *Remote Sens. Environ.*, 112(4), 1347–1364, doi:10.1016/j.rse.2007.05.020, 2008.
- Le Quéré, C., Raupach, M. R., Canadell, J. G., Al, G. M. et, Al, C. L. Q. et, Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viovy, N., Werf, G. R. van der and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, 2(12), 831–836, doi:10.1038/ngeo689, 2009.
- Rammig, A., Wiedermann, M., Donges, J. F., Babst, F., von Bloh, W., Frank, D., Thonicke, K. and Mahecha, M. D.: Coincidences of climate extremes and anomalous vegetation responses: comparing tree ring patterns to simulated productivity, *Biogeosciences*, 12(2), 373–385, doi:10.5194/bg-12-373-2015, 2015.
- Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y. and Field, C. B.: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Glob. Biogeochem. Cycles*, 11(4), 535–560, doi:10.1029/97GB02268, 1997.
- Raupach, Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S. and Schimmler, C. C.: Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, *Glob. Change Biol.*, 11(3), 378–397, 2005.
- Raynolds, M. K., Walker, D. A., Verbyla, D. and Munger, C. A.: Patterns of Change within a Tundra Landscape: 22-year Landsat NDVI Trends in an Area of the Northern Foothills of the Brooks Range, Alaska, *Arct. Antarct. Alp. Res.*, 45(2), 249–260, doi:10.1657/1938-4246-45.2.249, 2013.
- R Core Team: R: A language and environment for statistical computing., R Foundation for Statistical Computing, Vienna, Austria. [online] Available from: <http://www.R-project.org/>, 2014.

- Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W. and Ohlen, D. O.: Measuring Phenological Variability from Satellite Imagery, *J. Veg. Sci.*, 5(5), 703–714, doi:10.2307/3235884, 1994.
- Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogée, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., Grünwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. m., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.-F., Sanz, M.-J., Vesala, T. and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, *Glob. Change Biol.*, 13(3), 634–651, doi:10.1111/j.1365-2486.2006.01224.x, 2007.
- Richards, J. and Jia, X.: *Remote Sensing Digital Image Analysis*, 4th ed., Springer, Berlin., 2006.
- Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J. and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, *Glob. Change Biol.*, 18(2), 566–584, doi:10.1111/j.1365-2486.2011.02562.x, 2012.
- Richardson, A. D., Braswell, B. H., Hollinger, D. Y., Jenkins, J. P. and Ollinger, S. V.: Near-surface remote sensing of spatial and temporal variation in canopy phenology, *Ecol. Appl.*, 19(6), 1417–1428, doi:10.1890/08-2022.1, 2009.
- Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O. and Toomey, M.: Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, *Agric. For. Meteorol.*, 169, 156–173, doi:10.1016/j.agrformet.2012.09.012, 2013.
- Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F. and Heimann, M.: Interannual sea–air CO<sub>2</sub> flux variability from an



- observation-driven ocean mixed-layer scheme, *Biogeosciences*, 11(17), 4599–4613, doi:10.5194/bg-11-4599-2014, 2014.
- Rödenbeck, C., Houweling, S., Gloor, M. and Heimann, M.: CO<sub>2</sub> flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos Chem Phys*, 3(6), 1919–1964, doi:10.5194/acp-3-1919-2003, 2003.
- Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J. and Schaphoff, S.: Agricultural green and blue water consumption and its influence on the global water system, *Water Resour. Res.*, 44(9), W09405, doi:10.1029/2007wr006331, 2008.
- Running, S. W., Thornton, P. E., Nemani, R. and Glassy, J. M.: Global Terrestrial Gross and Net Primary Productivity from the Earth Observing System, in *Methods in Ecosystem Science*, edited by O. E. Sala, R. B. Jackson, H. A. Mooney, and R. W. Howarth, pp. 44–57, Springer New York. [online] Available from: [http://link.springer.com/chapter/10.1007/978-1-4612-1224-9\\_4](http://link.springer.com/chapter/10.1007/978-1-4612-1224-9_4) (Accessed 24 January 2015), 2000.
- Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M. and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, *Proc. Natl. Acad. Sci.*, 108(24), 9899–9904, doi:10.1073/pnas.1019576108, 2011.
- Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J. and Thonicke, K.: Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model, *Glob. Change Biol.*, n/a–n/a, doi:10.1111/gcb.12870, 2015.
- Saleska, S. R., Didan, K., Huete, A. R. and Rocha, H. R. da: Amazon Forests Green-Up During 2005 Drought, *Science*, 318(5850), 612–612, doi:10.1126/science.1146663, 2007.
- Samanta, A., Ganguly, S., Hashimoto, H., Devadiga, S., Vermote, E., Knyazikhin, Y., Nemani, R. R. and Myneni, R. B.: Amazon forests did not green-up during the 2005 drought, *Geophys. Res. Lett.*, 37(5), L05401, doi:10.1029/2009GL042154, 2010.
- Samanta, A., Ganguly, S., Vermote, E., Nemani, R. R. and Myneni, R. B.: Why Is Remote Sensing of Amazon Forest Greenness So Challenging?, *Earth Interact.*, 16(7), 1–14, doi:10.1175/2012EI440.1, 2012.

- Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L. and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, *Remote Sens. Environ.*, 83(1–2), 135–148, doi:10.1016/S0034-4257(02)00091-3, 2002.
- Schaphoff, S., Heyder, U., Ostberg, S., Gerten, D., Heinke, J. and Lucht, W.: Contribution of permafrost soils to the global carbon budget, *Environ. Res. Lett.*, 8(1), 014026, doi:10.1088/1748-9326/8/1/014026, 2013.
- Scheftic, W., Zeng, X., Broxton, P. and Brunke, M.: Intercomparison of Seven NDVI Products over the United States and Mexico, *Remote Sens.*, 6(2), 1057–1084, doi:10.3390/rs6021057, 2014.
- Schneider, U., Fuchs, T., Meyer-Christoffer, A. and Rudolf, B.: Global Precipitation Analysis Products of the GPCC, [online] Available from: [http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/KU/KU4/KU42/en/Reports\\_\\_Publications/GPCC\\_\\_intro\\_\\_products\\_\\_2011,templateId=raw,property=publicationFile.pdf/GPCC\\_intro\\_products\\_2011.pdf](http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/KU/KU4/KU42/en/Reports__Publications/GPCC__intro__products__2011,templateId=raw,property=publicationFile.pdf/GPCC_intro_products_2011.pdf), 2008.
- Schulze, E.-D.: Air Pollution and Forest Decline in a Spruce (*Picea abies*) Forest, *Science*, 244(4906), 776–783, doi:10.1126/science.244.4906.776, 1989.
- Schulze, E.-D.: Biological control of the terrestrial carbon sink, *Biogeosciences*, 3(2), 147–166, doi:10.5194/bg-3-147-2006, 2006.
- Schuur, E. A. G., Crummer, K. G., Vogel, J. G. and Mack, M. C.: Plant Species Composition and Productivity following Permafrost Thaw and Thermokarst in Alaskan Tundra, *Ecosystems*, 10(2), 280–292, doi:10.1007/s10021-007-9024-0, 2007.
- Schwartz, M. D. and Reed, B. C.: Surface phenology and satellite sensor-derived onset of greenness: An initial comparison, *Int. J. Remote Sens.*, 20(17), 3451–3457, doi:10.1080/014311699211499, 1999.
- Sellers, P. J.: Canopy reflectance, photosynthesis and transpiration, *Int. J. Remote Sens.*, 6(8), 1335–1372, doi:10.1080/01431168508948283, 1985.
- Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., Collatz, G. J., Denning, A. S., Mooney, H. A., Nobre, C. A., Sato, N., Field, C. B. and Henderson-Sellers, A.: Modeling the Exchanges of Energy, Water, and Carbon

- Between Continents and the Atmosphere, *Science*, 275(5299), 502–509, doi:10.1126/science.275.5299.502, 1997.
- Shanno, D. F.: Conditioning of quasi-Newton methods for function minimization, *Math. Comput.*, 24(111), 647–656, doi:10.1090/S0025-5718-1970-0274029-X, 1970.
- Sheffield, J. and Wood, E. F.: Global Trends and Variability in Soil Moisture and Drought Characteristics, 1950–2000, from Observation-Driven Simulations of the Terrestrial Hydrologic Cycle, *J. Clim.*, 21(3), 432–458, doi:10.1175/2007JCLI1822.1, 2008.
- Shur, Y. L. and Jorgenson, M. T.: Patterns of permafrost formation and degradation in relation to climate and ecosystems, *Permafr. Periglac. Process.*, 18(1), 7–19, 2007.
- Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S., Le Quéré, C., Smith, B., Zhu, Z. and Myneni, R.: Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades, *Biogeosciences Discuss*, 10(12), 20113–20177, doi:10.5194/bgd-10-20113-2013, 2013.
- Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C. and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), *Glob. Change Biol.*, 14(9), 2015–2039, doi:10.1111/j.1365-2486.2008.01626.x, 2008.
- Sitch, S., McGuire, A. D., Kimball, J., Gedney, N., Gamon, J., Engstrom, R., Wolf, A., Zhuang, Q., Clein, J. and McDonald, K. C.: Assessing the carbon balance of circumpolar arctic tundra using remote sensing and process modeling, *Ecol. Appl.*, 17(1), 213–234, doi:10.1890/1051-0761(2007)017[0213:ATCBOC]2.0.CO;2, 2007.
- Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Glob. Change Biol.*, (9), 161–185, 2003.
- Smith, B., Prentice, I. C. and Sykes, M. T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within

- European climate space, *Glob. Ecol. Biogeogr.*, 10(6), 621–637, doi:10.1046/j.1466-822X.2001.t01-1-00256.x, 2001.
- Smith, N. V., Saatchi, S. S. and Randerson, J. T.: Trends in high northern latitude soil freeze and thaw cycles from 1988 to 2002, *J. Geophys. Res. Atmospheres*, 109(D12), D12101, doi:10.1029/2003JD004472, 2004.
- Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O’Keefe, J. and Richardson, A. D.: Digital repeat photography for phenological research in forest ecosystems, *Agric. For. Meteorol.*, 152, 159–177, doi:10.1016/j.agrformet.2011.09.009, 2012.
- Specht, R.: Water Use by Perennial Evergreen Plant Communities in Australia and Papua New Guinea, *Aust. J. Bot.*, 20(3), 273–299, 1972.
- Specht, R. L.: Growth indices – Their rôle in understanding the growth, structure and distribution of Australian vegetation, *Oecologia*, 50(3), 347–356, doi:10.1007/BF00344975, 1981.
- Stöckli, R., Lawrence, D. M., Niu, G.-Y., Oleson, K. W., Thornton, P. E., Yang, Z.-L., Bonan, G. B., Denning, A. S. and Running, S. W.: Use of FLUXNET in the Community Land Model development, *J. Geophys. Res. Biogeosciences*, 113(G1), G01025, doi:10.1029/2007JG000562, 2008a.
- Stöckli, R., Rutishauser, T., Baker, I., Liniger, M. a and Denning, a S.: A global reanalysis of vegetation phenology, *J. Geophys. Res.*, 116(G3), G03020–G03020, doi:10.1029/2010jg001545, 2011.
- Stöckli, R., Rutishauser, T., Dragoni, D., O’Keefe, J., Thornton, P. E., Jolly, M., Lu, L. and Denning, A. S.: Remote sensing data assimilation for a prognostic phenology model, *J. Geophys. Res.*, 113(G4), G04021–G04021, 2008b.
- Stöckli, R., Vermote, E. F., Saleous, N., Simmon, R. and Herring, D.: The Blue Marble Next Generation - A true color earth dataset including seasonal dynamics from MODIS., [online] Available from: <http://visibleearth.nasa.gov/view.php?id=74393>, 2005.
- Stocks, B. J., Mason, J. A., Todd, J. B., Bosch, E. M., Wotton, M., Amiro, B. D., Flannigan, M. D., Hirsch, K. G., Logan, K. A., Martell, D. L. and Skinner, W. R.: Large forest fires in Canada, 1959-1997, *J. Geophys. Res.*, 107(D1), 8149–8149, doi:10.1029/2001jd000484, 2002.

- Strengers, B. J., Müller, C., Schaeffer, M., Haarsma, R. J., Severijns, C., Gerten, D., Schaphoff, S., van den Houdt, R. and Oostenrijk, R.: Assessing 20th century climate–vegetation feedbacks of land-use change and natural vegetation dynamics in a fully coupled vegetation–climate model, *Int. J. Climatol.*, 30(13), 2055–2065, doi:10.1002/joc.2132, 2010.
- Studer, S., Stöckli, R., Appenzeller, C. and Vidale, P. L.: A comparative study of satellite and ground-based phenology, *Int. J. Biometeorol.*, 51(5), 405–414, doi:10.1007/s00484-006-0080-5, 2007.
- Sturm, M., Racine, C. H. and Tape, K.: Climate change: Increasing shrub abundance in the Arctic, *Nature*, 411(6837), 546–547, doi:10.1038/35079180, 2001.
- Sugimoto, A., Yanagisawa, N., Naito, D., Fujita, N. and Maximov, T. C.: Importance of permafrost as a source of water for plants in east Siberian taiga, *Ecol. Res.*, 17, 493–503, 2002.
- Sulkava, M., Luyssaert, S., Rautio, P., Janssens, I. a and Hollmén, J.: Modeling the effects of varying data quality on trend detection in environmental monitoring, *Ecol. Inform.*, 2(2), 167–176, doi:10.1016/j.ecoinf.2007.03.008, 2007.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R. and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean pCO<sub>2</sub>, and net sea–air CO<sub>2</sub> flux over the global oceans, *Deep Sea Res. Part II Top. Stud. Oceanogr.*, 56(8–10), 554–577, doi:10.1016/j.dsr2.2008.12.009, 2009.
- Tao, X., Wang, D., Wu, D., Yan, B., Fan, W., Xu, X. and Yao, Y.: A model for instantaneous FAPAR retrieval: Theory and validation, in *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, vol. 1, pp. I–144–I–147., 2009.
- Tateishi, R. and Ebata, M.: Analysis of phenological change patterns using 1982–2000 Advanced Very High Resolution Radiometer (AVHRR) data, *Int. J. Remote Sens.*, 25(12), 2287–2300, doi:10.1080/01431160310001618455, 2004.

- Thonicke, K., Spessa, A., Prentice, I. C., Harrison, S. P., Dong, L. and Carmona-Moreno, C.: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model, *Biogeosciences*, 7(6), 1991–2011, 2010.
- Thonicke, K., Venevsky, S., Sitch, S. and Cramer, W.: The Role of Fire Disturbance for Global Vegetation Dynamics: Coupling Fire into a Dynamic Global Vegetation Model, *Glob. Ecol. Biogeogr.*, 10(6), 661–677, 2001.
- Thoning, K. W., Tans, P. P. and Komhyr, W. D.: Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985, *J. Geophys. Res. Atmospheres*, 94(D6), 8549–8565, doi:10.1029/JD094iD06p08549, 1989.
- Turner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R. and Schullius, C.: Carbon stock and density of northern boreal and temperate forests, *Glob. Ecol. Biogeogr.*, 23(3), 297–310, doi:10.1111/geb.12125, 2014.
- Le Toan, T., Quegan, S., Davidson, M. W. J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H. and Ulander, L.: The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle, *Remote Sens. Environ.*, 115(11), 2850–2860, doi:10.1016/j.rse.2011.03.020, 2011.
- Townshend, J., Carroll, M., Dimiceli, C., Sohlberg, R. a, Hansen, M. C. and DeFries, R. S.: Vegetation Continuous Fields MOD44B. 2000 Percent Tree Cover, Collection 5, [online] Available from: <http://glcf.umd.edu/data/vcf/> (Accessed 15 January 2013), 2011.
- Traore, A. K., Ciais, P., Vuichard, N., Poulter, B., Viovy, N., Guimberteau, M., Jung, M., Myneni, R. and Fisher, J. B.: Evaluation of the ORCHIDEE ecosystem model over Africa against 25 years of satellite-based water and carbon measurements, *J. Geophys. Res. Biogeosciences*, 119(8), 2014JG002638, doi:10.1002/2014JG002638, 2014.
- Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Liu, Q., Pak, B., Reichstein, M., Renzullo, L., Richardson, A. D., Roxburgh, S. H., Styles, J., Wang, Y. P., Briggs, P., Barrett, D. and Nikolova, S.: OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, *J. Geophys. Res. Biogeosciences*, 112(G2), G02027, doi:10.1029/2006JG000367, 2007.

- Tucker, C. J.: Red and Photographic Infrared Linear Combinations for Monitoring Vegetation, *Remote Sens. Environ.*, 150(8), 127–150, 1979.
- Tucker, C. J., Slayback, D. A., Pinzon, J. E., Los, S. O., Myneni, R. B. and Taylor, M. G.: Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999, *Int. J. Biometeorol.*, 45(4), 184–190, doi:10.1007/s00484-001-0109-8, 2001.
- Tucker, C., Pinzon, J., Brown, M., Slayback, D., Pak, E., Mahoney, R., Vermote, E. and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *Int. J. Remote Sens.*, 26(20), 4485–4498, doi:10.1080/01431160500168686, 2005.
- Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S. and Briggs, J. M.: Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across Three Temperate Zone Sites, *Remote Sens. Environ.*, 70(April 1998), 52–68, 1999.
- Tutubalina, O. V. and Rees, W. G.: Vegetation degradation in a permafrost region as seen from space: Noril'sk (1961–1999), *Cold Reg. Sci. Technol.*, 32(2–3), 191–203, doi:10.1016/S0165-232X(01)00049-0, 2001.
- Urban, M., Forkel, M., Eberle, J., Hüttich, C., Schmullius, C. and Herold, M.: Pan-Arctic Climate and Land Cover Trends Derived from Multi-Variate and Multi-Scale Analyses (1981–2012), *Remote Sens.*, 6(3), 2296–2316, doi:10.3390/rs6032296, 2014.
- Urban, M., Forkel, M., Schmullius, C., Hese, S., Hüttich, C. and Herold, M.: Identification of land surface temperature and albedo trends in AVHRR Pathfinder data from 1982 to 2005 for northern Siberia, *Int. J. Remote Sens.*, 34(12), 4491–4507, doi:10.1080/01431161.2013.779760, 2013.
- USGS: NASA Land Processes Distributed Active Archive Center (LP DAAC). MOD15A2 Leaf Area Index - Fraction of Photosynthetically Active Radiation 8-Day L4 Global 1km. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota., [online] Available from: [https://lpdaac.usgs.gov/products/modis\\_products\\_table/mod15a2](https://lpdaac.usgs.gov/products/modis_products_table/mod15a2), 2001.
- Veneables, W. N. and Ripley, B. D.: *Modern Applied Statistics with S*, Berlin. [online] Available from: <http://www.springer.com/mathematics/probability/book/978-0-387-95457-8> (Accessed 18 November 2014), 2002.

- Venevsky, S., Thonicke, K., Sitch, S. and Cramer, W.: Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study, *Glob. Change Biol.*, 8(10), 984–998, 2002.
- Verbesselt, J., Hyndman, R., Newnham, G. and Culvenor, D.: Detecting trend and seasonal changes in satellite image time series, *Remote Sens. Environ.*, 114(1), 106–115, doi:10.1016/j.rse.2009.08.014, 2010a.
- Verbesselt, J., Hyndman, R., Zeileis, A. and Culvenor, D.: Phenological change detection while accounting for abrupt and gradual trends in satellite image time series, *Remote Sens. Environ.*, 114(12), 2970–2980, doi:10.1016/j.rse.2010.08.003, 2010b.
- Verbesselt, J., Zeileis, A. and Herold, M.: Near real-time disturbance detection using satellite image time series, *Remote Sens. Environ.*, 123, 98–108, doi:10.1016/j.rse.2012.02.022, 2012.
- Verbyla, D.: The greening and browning of Alaska based on 1982–2003 satellite data, *Glob. Ecol. Biogeogr.*, 17(4), 547–555, 2008.
- Verbyla, D.: Browning boreal forests of western North America, *Environ. Res. Lett.*, 6(4), 041003, doi:10.1088/1748-9326/6/4/041003, 2011.
- Verstraeten, W. W., Veroustraete, F. and Feyen, J.: On temperature and water limitation of net ecosystem productivity: Implementation in the C-Fix model, *Ecol. Model.*, 199(1), 4–22, doi:10.1016/j.ecolmodel.2006.06.008, 2006.
- Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H. and Tilman, D. G.: Human alteration of the global nitrogen cycle: sources and consequences, *Ecol. Appl.*, 7(3), 737–750, doi:10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2, 1997.
- Vitousek, P. M. and Howarth, R. W.: Nitrogen limitation on land and in the sea: How can it occur?, *Biogeochemistry*, 13(2), 87–115, doi:10.1007/BF00002772, 1991.
- De Vries, W., Reinds, G. J., Gundersen, P. and Sterba, H.: The impact of nitrogen deposition on carbon sequestration in European forests and forest soils, *Glob. Change Biol.*, 12(7), 1151–1173, doi:10.1111/j.1365-2486.2006.01151.x, 2006.
- Wagner, W., Scipal, K., Pathe, C., Gerten, D., Lucht, W. and Rudolf, B.: Evaluation of the agreement between the first global remotely sensed soil moisture data with



- model and precipitation data, *J. Geophys. Res. Atmospheres*, 108(D19), 4611, doi:10.1029/2003JD003663, 2003.
- Walker, D. A., Leibman, M. O., Epstein, H. E., Forbes, B. C., Bhatt, U. S., Reynolds, M. K., Comiso, J. C., Gubarkov, A. A., Khomutov, A. V., Jia, G. J., Kaarlejärvi, E., Kaplan, J. O., Kumpula, T., Kuss, P., Matyshak, G., Moskalenko, N. G., Orekhov, P., Romanovsky, V. E., Ukraintseva, N. G. and Yu, Q.: Spatial and temporal patterns of greenness on the Yamal Peninsula, Russia: interactions of ecological and social factors affecting the Arctic normalized difference vegetation index, *Environ. Res. Lett.*, 4(4), 045004, doi:10.1088/1748-9326/4/4/045004, 2009.
- Walker, J. J., de Beurs, K. M. and Wynne, R. H.: Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data, *Remote Sens. Environ.*, 144, 85–97, doi:10.1016/j.rse.2014.01.007, 2014.
- Walter, H.: Ergänzende Betrachtungen zu der im Klimadiagramm-Weltatlas verwendeten Klimadarstellung, *Erdkunde*, 24(2), 145–149, 1970.
- Walter, H. and Lieth, H.: *Klimadiagramm-Weltatlas*, Fischer, Jena. [online] Available from: <http://tinyurl.sfx.mpg.de/u2za>, 1960.
- Walter-Shea, E. A., Blad, B. L., Mesarch, M. A., Hays, C. J., Deering, D. W. and Eck, T. F.: Absorbed photosynthetically active radiation and sun-view geometry effects on remote sensing relationships, *Remote Sens. Rev.*, 17(1-4), 89–102, doi:10.1080/02757259809532365, 1998.
- Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O. and Bairlein, F.: Ecological responses to recent climate change, *Nature*, 416(6879), 389–395, doi:10.1038/416389a, 2002.
- Wang, D., Morton, D., Masek, J., Wu, A., Nagol, J., Xiong, X., Levy, R., Vermote, E. and Wolfe, R.: Impact of sensor degradation on the MODIS NDVI time series, *Remote Sens. Environ.*, 119, 55–61, doi:10.1016/j.rse.2011.12.001, 2012.
- Wang, X., Piao, S., Ciais, P., Li, J., Friedlingstein, P., Koven, C. and Chen, A.: Spring temperature change and its implication in the change of vegetation growth in North America from 1982 to 2006, *Proc. Natl. Acad. Sci. U. S. A.*, 108(4), 1240–1245, doi:10.1073/pnas.1014425108, 2011.

- Wang, Y.-P., Trudinger, C. M. and Enting, I. G.: A review of applications of model–data fusion to studies of terrestrial carbon fluxes at different scales, *Agric. For. Meteorol.*, 149(11), 1829–1842, doi:10.1016/j.agrformet.2009.07.009, 2009.
- Wang, Y., Roderick, M. L., Shen, Y. and Sun, F.: Attribution of satellite-observed vegetation trends in a hyper-arid region of the Heihe River basin, Western China, *Hydrol Earth Syst Sci*, 18(9), 3499–3509, doi:10.5194/hess-18-3499-2014, 2014.
- Watanabe, F., Uchino, O., Joo, Y., Aono, M., Higashijima, K., Hirano, Y., Tsuboi, K. and Suda, K.: Interannual variation of growth rate of atmospheric carbon dioxide concentration observed at the JMA’s three monitoring stations: Large increase in concentration of atmospheric carbon dioxide in 1998, *J. Meteorol. Soc. Jpn.*, 78(5), 673, 2000.
- Van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y. and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), *Atmos Chem Phys*, 10(23), 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.
- Wermelinger, B.: Ecology and management of the spruce bark beetle *Ips typographus*—a review of recent research, *For. Ecol. Manag.*, 202(1–3), 67–82, doi:10.1016/j.foreco.2004.07.018, 2004.
- White, K., Pontius, J. and Schaberg, P.: Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty, *Remote Sens. Environ.*, 148, 97–107, doi:10.1016/j.rse.2014.03.017, 2014.
- White, M. A., Beurs, K. M. de, Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O’Keefe, J., Zhang, G., Nemani, R. R., van Leeuwen, W. J. D., Brown, J. F., Wit, A. de, Schaepman, M., Lin, X., Dettinger, M., Bailey, A. S., Kimball, J., Schwartz, M. D., Baldocchi, D. D., Lee, J. T. and Lauenroth, W. K.: Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006, *Glob. Change Biol.*, 15(10), 2335–2359, 2009.
- White, M. A., Thornton, P. E. and Running, W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability, *Glob. Biogeochem. Cycles*, 11(2), 217–234, 1997.

- Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M. and Wang, Y. P.: Improving land surface models with FLUXNET data, *Biogeosciences*, 6(7), 1341–1359, doi:10.5194/bg-6-1341-2009, 2009.
- Wolkovich, E. M. and Cleland, E. E.: The phenology of plant invasions: a community ecology perspective, *Front. Ecol. Environ.*, 9(5), 287–294, doi:10.1890/100033, 2010.
- Wolkovich, E. M., Cook, B. I., Allen, J. M., Crimmins, T. M., Betancourt, J. L., Travers, S. E., Pau, S., Regetz, J., Davies, T. J., Kraft, N. J. B., Ault, T. R., Bolmgren, K., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Salamin, N., Schwartz, M. D. and Cleland, E. E.: Warming experiments underpredict plant phenological responses to climate change, *Nature*, 485(7399), 494–497, doi:10.1038/nature11014, 2012.
- Worthy, D. E. J.: Canadian Baseline Program, Meteorological Service of Canada, Downsview, Ontario, Canada., 2003.
- Wright, S. J.: Phenological Responses to Seasonality in Tropical Forest Plants, in *Tropical Forest Plant Ecophysiology*, edited by S. S. Mulkey, R. L. Chazdon, and A. P. Smith, pp. 440–460, Springer US. [online] Available from: [http://link.springer.com/chapter/10.1007/978-1-4613-1163-8\\_15](http://link.springer.com/chapter/10.1007/978-1-4613-1163-8_15) (Accessed 26 February 2014), 1996.
- Wu, C., Gonsamo, A., Gough, C. M., Chen, J. M. and Xu, S.: Modeling growing season phenology in North American forests using seasonal mean vegetation indices from MODIS, *Remote Sens. Environ.*, 147, 79–88, doi:10.1016/j.rse.2014.03.001, 2014.
- Wutzler, T. and Carvalhais, N.: Balancing multiple constraints in model-data integration: Weights and the parameter-block approach, *J. Geophys. Res. Biogeosciences*, 2014JG002650, doi:10.1002/2014JG002650, 2014.
- Wu, Z., Schneider, E. K., Kirtman, B. P., Sarachik, E. S., Huang, N. E. and Tucker, C. J.: The modulated annual cycle: an alternative reference frame for climate anomalies, *Clim. Dyn.*, 31(7-8), 823–841, doi:10.1007/s00382-008-0437-z, 2008.
- Xiao, X., Zhang, J., Yan, H., Wu, W. and Biradar, C.: Land Surface Phenology, in *Phenology of Ecosystem Processes*, edited by A. Noormets, pp. 247–270, Springer New York. [online] Available from: [http://link.springer.com/chapter/10.1007/978-1-4419-0026-5\\_11](http://link.springer.com/chapter/10.1007/978-1-4419-0026-5_11) (Accessed 23 July 2014), 2009.

- Xu, L., Myneni, R. B., Chapin Iii, F. S., Callaghan, T. V., Pinzon, J. E., Tucker, C. J., Zhu, Z., Bi, J., Ciais, P., Tømmervik, H., Euskirchen, E. S., Forbes, B. C., Piao, S. L., Anderson, B. T., Ganguly, S., Nemani, R. R., Goetz, S. J., Beck, P. S. A., Bunn, A. G., Cao, C. and Stroeve, J. C.: Temperature and vegetation seasonality diminishment over northern lands, *Nat. Clim. Change*, 3(6), 581–586, doi:10.1038/nclimate1836, 2013.
- Yuan, W., Zhou, G., Wang, Y., Han, X. and Wang, Y.: Simulating phenological characteristics of two dominant grass species in a semi-arid steppe ecosystem, *Ecol. Res.*, 22(5), 784–791, doi:10.1007/s11284-006-0318-z, 2007.
- Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warlind, D., Weng, E., Parton, W., Iversen, C. M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P. J., Prentice, I. C., Oren, R. and Norby, R. J.: Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO<sub>2</sub> Enrichment studies, *New Phytol.*, 202(3), 803–822, doi:10.1111/nph.12697, 2014.
- Zaehle, S., Sitch, S., Smith, B. and Hatterman, F.: Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics, *Glob. Biogeochem. Cycles*, 19(3), doi:10.1029/2004gb002395, 2005.
- Zeileis, A.: *ineq: Measuring Inequality, Concentration, and Poverty*. [online] Available from: <http://CRAN.R-project.org/package=ineq>, 2014.
- Zeileis, A., Kleiber, C., Krämer, W. and Hornik, K.: Testing and dating of structural changes in practice, *Comput. Stat. Data Anal.*, 44(1-2), 109–123, doi:10.1016/s0167-9473(03)00030-6, 2003.
- Zeng, F.-W., Collatz, G., Pinzon, J. and Ivanoff, A.: Evaluating and Quantifying the Climate-Driven Interannual Variability in Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) at Global Scales, *Remote Sens.*, 5(8), 3918–3950, doi:10.3390/rs5083918, 2013.
- Zeng, H., Jia, G. and Epstein, H.: Recent changes in phenology over the northern high latitudes detected from multi-satellite data, *Environ. Res. Lett.*, 6(4), 045508–045508, doi:10.1088/1748-9326/6/4/045508, 2011.

- Zeng, N., Zhao, F., Collatz, G. J., Kalnay, E., Salawitch, R. J., West, T. O. and Guanter, L.: Agricultural Green Revolution as a driver of increasing atmospheric CO<sub>2</sub> seasonal amplitude, *Nature*, 515(7527), 394–397, doi:10.1038/nature13893, 2014.
- Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, *Rev. Geophys.*, 43(4), 2005.
- Zhang, X., Friedl, M. A. and Schaaf, C. B.: Sensitivity of vegetation phenology detection to the temporal resolution of satellite data, *Int. J. Remote Sens.*, 30(8), 2061–2074, doi:10.1080/01431160802549237, 2009.
- Zhang, X., Friedl, M. A., Schaaf, C. B. and Strahler, A. H.: Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data, *Glob. Change Biol.*, 10(7), 1133–1145, doi:10.1111/j.1529-8817.2003.00784.x, 2004.
- Zhang, X., Tan, B. and Yu, Y.: Interannual variations and trends in global land surface phenology derived from enhanced vegetation index during 1982–2010, *Int. J. Biometeorol.*, 58(4), 547–564, doi:10.1007/s00484-014-0802-z, 2014.
- Zhang, X., Tarpley, D. and Sullivan, J. T.: Diverse responses of vegetation phenology to a warming climate, *Geophys. Res. Lett.*, 34(19), L19405, doi:10.1029/2007GL031447, 2007.
- Zhou, L., Tian, Y., Myneni, R. B., Ciais, P., Saatchi, S., Liu, Y. Y., Piao, S., Chen, H., Vermote, E. F., Song, C. and Hwang, T.: Widespread decline of Congo rainforest greenness in the past decade, *Nature*, 509(7498), 86–90, doi:10.1038/nature13265, 2014.
- Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. and Myneni, R.: Global Data Sets of Vegetation Leaf Area Index (LAI)<sub>3g</sub> and Fraction of Photosynthetically Active Radiation (FPAR)<sub>3g</sub> Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI<sub>3g</sub>) for the Period 1981 to 2011, *Remote Sens.*, 5(2), 927–948, doi:10.3390/rs5020927, 2013.

## Acknowledgement / Danksagung / Poděkování

I thank Nuno Carvalhais for being a great advisor and a critical companion during all challenges. You were motivating me to leave no stone unturned. My thanks also go to Markus Reichstein, Christiane Schullius and Kirsten Thonicke for their supervision, for critical questions and comments, and for providing me the basis to grow in science.

Mirco Migliavacca has been an inexhaustible source of knowledge and literature.

Christian Beer, Martin Heimann, Ralph Keeling, Miguel Mahecha, Christopher Neigh, Martin Thurner, Jan Verbesselt, and Sönke Zaehle provided ideas, data, and many critical feedbacks to our joint manuscripts.

Sibyll Schaphoff, Werner von Bloh and Christian Rödenbeck have been indispensable supporters in modelling, code merging and debugging.

Ulrich Weber has been an essential express supplier of global datasets (and drinks).

Martin Jung, Enrio Tomelleri and Maarten Braakhekke helped me with data, analyses and feedbacks during different stages of this research.

Martin Thurner, Altug Ekici and Sonja Kaiser were great office mates and friends.

Marcel Urban and Lyudmila Lebedeva gave me the opportunity to contribute with my ideas to their work.

Franziska Schrodtt and Julia Marshall provided important feedback on language and on the introduction and conclusions of this thesis.

I am grateful to all friends and colleagues for a panoply of support and conversations: Bernhard Ahrens, Steffen Beer, Sven Boese, Jonas Eberle, Hans-Jörg Fischer, Dorothea Frank, Eberhard Fritz, Myroslava Khomik, Sujan Koirala, Karl Kübler, Gitta Lasslop, Lee Miller, Jörg Reith, Maik Renner, Boris Sakschewski, Silvana Schott, Gregor Schürmann, Kerstin Sickel, Bertram Smolny, Mikhail Urbazaev, Birgitta Wiehl, and Jakob Zscheischler.

Vielen Dank meinen Eltern, Ursula und Frank Forkel, für die vorbehaltlose Unterstützung.

Mnohokrát děkuji Marii a Liboru Zajíčkovi za podporu s Helenkou a za všechno.

Helenka, du hast mir die Zeit gegeben über meine Arbeit nachzudenken wann immer ich dich im Kinderwagen durch den Wald geschoben habe. Danke für dein Lachen!

Lenka, děkuji za tvoje láska. Mehr kann ich nicht in Worte fassen.

Thank you all for having contributed to my research.

## **Author contributions to the manuscripts**

### **Erklärung zu den Eigenanteilen des Promovenden an den Publikationen und Zweitpublikationsrechten bei einer kumulativen Dissertation**

Für alle in dieser kumulativen Dissertation verwendeten Manuskripte liegen die notwendigen Genehmigungen der Verlage („Reprint permissions“) für die Zweitpublikation vor.

Die Ko-Autoren der in dieser kumulativen Dissertation verwendeten Manuskripte sind sowohl über die Nutzung, als auch über die oben angegebenen Eigenanteile informiert und stimmen dem zu.

Die Anteile der Ko-Autoren an den Publikationen sind nachfolgend aufgeführt.

Ich bin mit der Abfassung der Dissertation als publikationsbasiert, d.h. kumulativ, einverstanden und bestätige die vorstehenden Angaben. Eine entsprechend begründete Befürwortung mit Angabe des wissenschaftlichen Anteils des Doktoranden/der Doktorandin an den verwendeten Publikationen werde ich parallel an den Rat der Fakultät der Chemisch-Geowissenschaftlichen Fakultät richten.

Name des Betreuers

Ort

Datum

Unterschrift

Publication 1:

Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M., Neigh, C. and Reichstein, M.: Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology, *Remote Sensing*, 5(5), 2113–2144, doi:10.3390/rs5052113, 2013.

Author initials	MF	NC	JV	MM	CN	MR
Conception of research	X	X	X			X
Planning of experiments	X	X		X		
Method development and code implementation	X		X	X		
Calculations	X					
Providing and interpretation of data					X	
Analysis and interpretation of results	X	X	X	X	X	X
Writing of manuscript	X	X				
Comments on manuscript		X	X	X	X	X

Publication 2:

Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

Author initials	MF	NC	SS	WvB	MM	MT	KT
Conception of research	X	X					X
Planning of experiments	X	X					
Method development and code implementation	X		X	X			
Calculations	X		X	X			
Providing and interpretation of data						X	
Analysis and interpretation of results	X	X	X		X	X	X
Writing of manuscript	X						
Comments on manuscript		X	X	X	X	X	X



---

Manuscript 3:  
Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U. and  
Carvalhais, N.: Co-dominant water control on global inter-annual variability and trends in land  
surface phenology and greenness, submitted to *Global Change Biology*, 2014-12-16.

---

Author initials	MF	MM	KT	MR	SS	UW	NC
Conception of research	X	X	X	X			X
Planning of experiments	X						X
Method development and code implementation	X	X			X		
Calculations	X						
Providing and processing of data	X					X	
Analysis and interpretation of results	X	X	X	X	X		X
Writing of manuscript	X		X	X			
Comments on manuscript		X	X		X		X

---

Manuscript 4:  
Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R. F., Heimann, M., Thonicke, K., Zaehle, S.  
and Reichstein, M.: Enhanced seasonal CO<sub>2</sub> exchange caused by amplification of plant  
productivity in the northern biosphere, submitted to *Nature*, 2015-01-30.

---

Author initials	MF	NC	CR	RK	MH	KT	SZ	MR
Conception of research	X	X						X
Planning of experiments	X	X	X					
Calculations	X		X					
Providing and processing of data	X		X	X				
Analysis and interpretation of results	X	X	X	X	X	X	X	X
Writing of manuscript	X	X						X
Comments on manuscript			X	X	X	X	X	

---

## Curriculum vitae

Name **Matthias Forkel**  
Birth 12th June 1985 in Zittau, Germany  
Family status married, one daughter  
Contacts Max Planck Institute for Biogeochemistry  
Hans-Knöll-Str. 10  
07745 Jena  
Germany  
  
mforkel@bgc-jena.mpg.de, matthias@m-forkel.de  
+0049 (0) 3641 576283

### Professional experiences and internships

- 01/2011-01/2015 *Doctoral candidate* at Max Planck Institute for Biogeochemistry Jena, Friedrich Schiller University Jena and International Max Planck Research School for Global Biogeochemical Cycles  
Advising committee: Dr. Nuno Carvalhais, Prof. Dr. Markus Reichstein, Prof. Dr. Christiane Schmullius and Dr. Kirsten Thonicke  
Performed research on the integration of remote sensing data in global ecosystem models  
Co-ordinated the working package “integration data” in the EU FP7 project Carbones (30-year re-analysis of CARBON fluxES and pools over Europe and the globe)
- 06/2010-11/2010 *Visiting MSc candidate* at Potsdam Institute for Climate Impact Research
- 08/2009-11/2010 *Scientific assistant* at Department for Earth Observation, University Jena  
Organized and led a field campaign of hyperspectral measurements as part of the ESA-funded EO-LDAS project (Earth Observation Land Data Assimilation System)
- 03/2010-04/2010 *Internship* at Max Planck Institute for Biogeochemistry  
Analyses on the integration of eddy covariance data in ecosystem models
- 03/2009-04/2009 *Internship* at Potsdam Institute for Climate Impact Research  
Analyses on the role of fire and vegetation in the Amazon
- 11/2004-05/2005 *Freelancer* at Klett-Perthes Publisher  
Authored educational material about climate for an online service

### Education

- 12/2010 *Master of Science* (Geoinformatics), University Jena  
Grade: 1.0 (ECTS: A), Thesis: 1.0  
Thesis: Permafrost-Fire Interactions in the Circumpolar Boreal Zone - An Analysis using Earth Observation Products and Simulations with a Dynamic Global Vegetation Model based on Global Climate Projections  
Advisors: Prof. Dr. Christiane Schmullius and Dr. Kirsten Thonicke
- 10/2008-12/2010 *Study of Geoinformatics* (M. Sc.), University Jena  
Majors: remote sensing, hydrological modelling, GIS

- 09/2008 *Bachelor of Science (Geography), University Jena*  
 Grade: 1.3 (ECTS: A), Thesis: 1.0  
 Thesis: Raum-zeitliche Analyse atmosphärischen Kohlenstoffs in Sibirien  
 mit ENVISAT-SCIAMACHY-Daten  
 Advisors: Prof. Dr. Christiane Schmullius and Roman Gerlach
- 10/2005-09/2008 *Study of Geography (B. Sc.), University Jena*  
 Major subjects: remote sensing, geoinformatics, soil science, geoecology
- 06/2004 *Abitur, Richard-von-Schlieben Gymnasium Zittau*  
 Grade: 1.3

### **Social Civil Service**

- 09/2004-06/2004 *Zivildienst at Lisa-Tetzner-Schule für Geistigbehinderte, Zittau*  
 Cared for mentally disabled children

### **Teaching, outreach activities and skills**

- 2012-2014 *Supporting several public outreach activities at Max Planck Institute Jena, e. g.*  
 Long Night of Sciences ("Why is the Earth green?", satellites and global  
 vegetation models), school project days ("photosynthesis"), lectures at  
 schools ("global carbon cycle")
- 2010-2015 *Teaching courses on "Analysis of spatial data in R" at the Department for*  
 Earth Observation, University Jena and within International Max Planck  
 Research School for Global Biogeochemical Cycles
- 2012 *Teaching a course on "GPS and GIS in ecology", Institute for Ecology,*  
 University Jena
- 2008-2010 *Tutoring a lecture on "statistics for geographers", Institute for Geography,*  
 University Jena
- 2002-2008 *Development of a private educational website on climate geography,*  
[www.klima-der-erde.de](http://www.klima-der-erde.de)
- Computational *Operation systems: Window, Unix*  
 skills *Programming: C, Java*  
*Data analysis: R, developer of the "greenbrown" package*  
*Web: HTML, PHP, MySQL*  
*Geographic Information Systems: ArcGIS, SAGA*  
*Remote sensing: ENVI, eCognition, PCI Geomatica*
- Language skills *German – native language*  
*English – fluent, writing and orally*  
*French – basic, writing and orally*  
*Czech – basic, orally*

## **Selbstständigkeitserklärung**

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe.

Jena, 1. Juni 2015

---

Matthias Forkel

