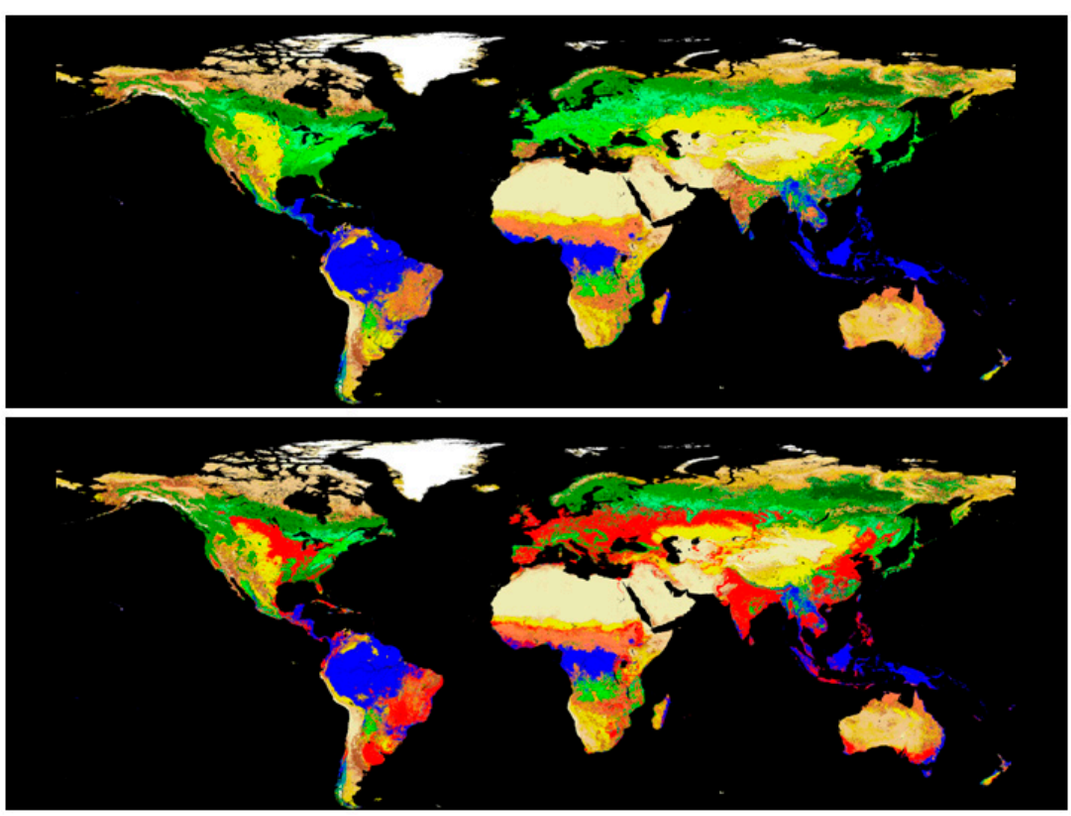
Max-Planck-Institut für Biogeochemie

MPI-BGC Jena

TECHNICAL REPORTS 11



Uncertainties of terrestrial carbon cycle modelling: Studies on gross carbon uptake of Europe

by
Martin Jung

ISSN 1615-7400

MAX-PLANCK-GESELLSCHAFT

Technical Reports - Max-Planck-Institut für Biogeochemie 11, 2008

Max-Planck-Institut für Biogeochemie
P.O.Box 10 01 64
07701 Jena/Germany
phone: +49 3641 576-0
fax: + 49 3641 577300
http://www.bgc-jena.mpg.de

Uncertainties of terrestrial carbon cycle modelling:

Studies on gross carbon uptake of Europe

Dissertation

Zur Erlangung des Doktergrades der Naturwissenschaften im Department für Geowissenschaften der Universität Hamburg

vorgelegt von

Martin Jung

aus

Erfurt

Hamburg

2008

Als Dissertation angenommen vom Department Geowissenschaften der Universität Hamburg

Auf Grund der Gutachten von Prof. Dr. Martin Claussen
und Dr. Galina Churkina

Hamburg, den 21.12.2007
Professor Dr. Kay-Christian Emeis
Leiter des Departments für Geowissenschaften

Acknowledgements

I want to thank my advisors Galina Churkina, Martin Herold, Martin Heimann, and Martin Claussen for their continuous support and stimulation. I want to thank a number of nice collaborators for their contributions and eye opening discussions, most importantly Markus Reichstein, Nadine Gobron, Sebastiaan Luyssaert, Sönke Zaehle, Guerric Le Maire, Bernard Pinty, Michel Verstraete, Mona Vetter, Philippe Ciais, and Alberte Bondeau. A particular thanks to Nadine again for her successful fight against the EU bureaucracy dragon, and managing that I could spend four months at the JRC in Ispra, which was too good. Financial support by the IMPRS exchange program and the DAAD made this visit finally possible.

Thanks to Sandy Harrison who gave me the opportunity to start a PhD at the MPI-BGC even though it didn't work out between us. In this respect, I am very grateful to Martin Heimann who opened the opportunity to continue my PhD with a new orientation and who almost successfully converted me into a carbon cycle modeller within an hour meeting. Thanks to Petra Bauer and Detlef Schulze for not introducing fixed working hours for PhD students at the MPI-BGC.

In fact, I should thank Tante Beate with whom I decided to study Geography on some sunny morning at the Baltic Sea between written and oral A-level exams in 1997. I would probably have studied something (even more) useless. However, if I acknowledge Tante Beate, I should also acknowledge my former sports teacher Herrn Altstädt who was cool and I was convinced that he would get the geography course – the only reason why I chose it (but it was Frau Dommes, not so nice at all, and I would love to tell her that the Pleistocene did not start 1.5 million years ago).

There are friends who have been and are very important to me but I don't feel like listing them here. My parents are the best!

Table of Contents

Figures and tables 5

Alphabetic list of frequently used acronyms 9

Abstract 11

1 Introduction 13
1.1 Motivation and Objectives 13
1.2 Outline of the thesis 15
2 Exploiting synergies of global land cover products for carbon cycle modelling 17
Abstract 17
2.1 Introduction 18
2.2 Aims and objectives 19
2.3 Overview of existing global land cover data sets 20
2.3.1 General principle 20
2.3.2 NOAA-AVHRR (GLCC) 21
2.3.3 SPOT-VEGETATION (GLC 2000) 22
2.3.4 TERRA-MODIS 22
2.3.5 Advantages and shortcomings of land cover mapping approaches 23
2.3.6 Validation of global land cover products 25
2.4 (Dis)Agreement of GLCC, GLC2000 and MODIS land cover products 27
2.4.1 Method 27
2.4.2 Result 29
2.4.3 Discussion 30
2.5 Land cover data fusion – exploring synergies between land cover products 31
2.5.1 General principle 31
2.5.2 Definition of the target legend 32
2.5.3 Selection and pre-processing of input data sets 32
2.5.4 Definition of affinity scores 34
2.5.5 Calculation of SYNMAP 35
2.6 SYNMAP evaluation 39
2.7 Limitations and advantages of SYNMAP 43
2.8 The way forward from a user's perspective 45
2.9 Summary and conclusions 46
5.10 Appendix – Affinity scores for life forms, leaf types and leaf longevities 49

3 Uncertainties of modelling GPP over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models 55
Abstract 55
3.1 Introduction 56
3.2 Biosphere models and driver data set options 57
3.2.1 Terrestrial biosphere models 57
3.2.2 Meteorological and land cover forcings 58
3.3 Experimental design 59

3.3.1 Modelling strategy 59
3.3.2 Quantification of effects 60
3.3.3 Decomposing GPP into absorbed photosynthetic active radiation and radiation use efficiency 61
3.3.4 Investigating the models' response to meteorology 62
3.4 Results and discussion 63
3.4.1 Order of effects 63
3.4.2 Land cover 66
3.4.3 Spatial resolution of the land cover map 67
3.4.4 Daily meteorology 68
3.4.5 Biosphere models 70
3.5 Conclusions 76
3.6 Auxiliary material 77

4 Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe 80
Abstract 80
4.1 Introduction 81
4.2. Materials and Methods 82
2.1 Site data 82
4.2.2 Model simulations 83
4.2.3 Decomposing GPP into APAR and RUE 84
4.3 Results and Discussion 86
4.3.1 Gross Primary Productivity 86
4.3.2 Leaf Area Index 90
4.3.3 Decomposing GPP into APAR and RUE 92
4.4 Conclusions 95

5 Diagnostic assessment of European gross primary production 96
Abstract 96
5.1 Introduction 97
5.2 Relating the cumulative growing season FAPAR to gross carbon uptake 100
5.2.1 Materials and Methods 100
5.2.2 Results and Discussion 105
5.3 Up-scaling GPP to Europe and corroboration with independent models 111
5.3.1 Materials and Methods 111
5.3.2 Results and Discussion 113
5.4 Conclusions 122
6 Summary, conclusions, and final remarks 123
6.1 What are the major sources of uncertainties of process-oriented modelling of GPP for Europe? 123
6.2 How realistic are GPP simulations from process-oriented models for Europe? 124
6.3 What is the GPP of Europe? 125
6.4 Remarks on evaluations of global terrestrial carbon cycle models 126
6.5 Towards reducing uncertainties of global terrestrial carbon cycle models 129
References 131

Figures and tables

Figure 1-1: Maps of agreement and disagreement between land cover products. 29

Figure 2-2: Principle of the data fusion method. 36

Figure 2-3: The SYNMAP data set. 39

Figure 2-4: Overall consistency between GLCC, GLC2000, MODIS and SYNMAP based on major life forms (SIMPLE legend) (a) and leaf attributes (b). 40

Figure 2-5: Average class specific consistencies for GLCC, GLC2000, MODIS and SYNMAP derived from intermap comparison (filled markers) and 'ground truth' class accuracies derived from published confusion matrices for GLCC, GLC2000 and MODIS based on validation data (not filled markers). 42

Figure 3-1: Simulation strategy to assess model performance differences due to the choice of the driver data set and carbon cycle model. 60

Figure 3-2: Difference maps of mean European GPP 1981-2000 for alternative realisations (AR-REF). 64

Figure 3-3: Effect of alternative realisations on the interannual variation of GPP. 65

Figure 3-4: Effects of different model set-ups (alternative realisations) on the magnitude, spatial, and temporal pattern on GPP simulations over Europe. 66

Figure 3-5: Coefficient of variation (standard deviation divided by mean, in %) of GPP, APAR, and RUE for Biome-BGC, LPJ, and Orchidee (1981-2000). 72

Figure 3-6: Correlation and sensitivity (slope of regression line) of relative GPP variations to the first principal component of mean JJA meteorology. 73

Figure 3-7: The fraction of interannual variance that is not explained by the correlation R^2 between LPJ and Orchidee for each pixel. 77

Figure 3-8: Difference of mean annual (1981-2000) meteorological variables of ECMWF and REMO (ECMWF-REMO). 77

Figure 3-9: The fraction of variance that is not explained by the correlation R^2 ($1-R^2$) between meteorological forcing fields between ECMWF and REMO for each pixel. 78

Figure 3-10: Fractions of the most important vegetation types used as input and mean maximum LAI and annual GPP (1981-2000) as simulated by Biome-BGC, LPJ, and Orchidee. 78

Figure 3-11: Comparison of the variation of APAR for different models and an independent estimate calculated from the SeaWiFS FAPAR product. 79

Figure 3-12: Maps of Pearson's correlation coefficient between GPP and APAR and RUE (1981-2000). 79

Figure 4-1: Spatial distribution of GPP and LAI measurements. 82

Figure 4-2: Top panel: eddy covariance flux separated (filled markers) and modelled (open markers) GPP along the mean annual temperature gradient across Europe. Bottom panel: difference between modelled and eddy covariance flux separated GPP along mean annual temperature (MAT, 1981-2000 mean based on the REMO data set). 87

Figure 4-3: Top panel: eddy covariance flux separated (filled markers) and modelled (open markers) GPP along a gradient of water availability for sites south of 52° latitude. Bottom panel: difference between modelled and eddy covariance flux separated GPP along the gradient of water availability. 89

Figure 4-4: Top panel: observed (filled markers) and modelled (open markers) maximum fAPAR along the mean annual temperature gradient across Europe. Bottom panel: difference between modelled and observed fAPAR along MAT. 91

Figure 4-5: Site (filled markers) and modelled (open markers) trends of APAR and RUE along the mean annual temperature gradient for boreal and temperate coniferous forests. 92

Figure 4-6: Coefficient of variation (standard deviation/mean) of APAR and RUE for boreal and temperate coniferous forests based on site and modelled data. 94

Figure 5-1: Map of Europe with CarboEurope sites used in this study. 102

Figure 5-2: Illustration of the algorithm to calculate the cumulative FAPAR of the growing season. 104

Figure 5-3: Map of mean growing season length (1998-2002) based on the proposed algorithm to calculate the cumulative FAPAR of the growing season. 105

Figure 5-4: Scatter plots of the cumulative growing season FAPAR and GPP for (a) all data points, (b) stratified by ecosystem types. 106

Figure 5-5: Maps of the 2000-2002 mean GPP from LPJmL, ANN, MOD17+, FPA, and FPA+LC. 114

Figure 5-6: Intercomparison of spatial patterns of JRC-FAPAR, MODIS-FAPAR, and GPP estimates from FPA+LC, ANN, and MOD17+ for (a) the 2000-2002 mean, and (b) the 2003 anomaly. 115

Figure 5-7: Maps of the 2003 anomaly of GPP from LPJmL, ANN, FPA, and FPA+LC. 119

Figure 5-8: (a) defined regions of the European domain, (b) time series of GPP for four major regions as predicted by LPJmL, ANN, FPA, and FPA+LC. 120

Table 2-1: Global land cover products with 1km spatial resolution used in this study. 21

Table 2-2: Table with translation between the SIMPLE-legend and the IGBP-DIScover (for GLCC and MODIS) and LCCS (GLC2000) legends. 28

Table 2-3: SYNMAP legend defined by dominant life form assemblage and tree leaf attributes. 33

Table 2-4: Definition of affinity scores according to semantic rules. The example uses the IGBP-Discover class ‘Woody savanna’. 35

Table 2-5: Calculation example for the best estimate of life form assemblage along the pixel vector of the land cover data sets. 37

Table 2-6: Calculation example for leaf type. 38

Table 2-7: Decision matrix for leaf type (below diagonal) and longevity (above diagonal) in case two leaf classes receive the same score. 68

Table 3-1: Total GPP of European domain as simulated by different model set-ups (1981-2000 mean). 64

Table 3-2: Result of the principal component analysis (PCA) of the meteorological input data. 73

Table 4-1: Relative RMSE and mean eddy covariance flux separated and modelled GPP, stratified by forest ecosystem type. 86

Table 4-2: Trends of APAR and RUE along MAT for boreal and temperate evergreen needleleaf forests. 93

Table 5-1: Statistics on the relationship between the cumulative FAPAR of the growing season and annual sums of GPP for different groups of ecosystems. 106

Table 5-2: Compilation of RMSE and R^2 values for data-driven GPP models from multi-site studies using eddy covariance GPP estimates. 108

Table 5-3: Pearson's correlation (R2) between GPP and the index of water availability (IWA), mean annual temperature (MAT), the annual sum of incoming shortwave radiation (RAD), the cumulative FAPAR of the growing season (cum GSL FAPAR), and absorbed radiation (ARAD). 111

Table 5-4: Matrix of Pearson's correlation coefficients of spatial GPP patterns as predicted by LPJmL, MOD17+, ANN, FPA, and FPA+LC. 113

Table 5-5: Total GPP flux of the 2000 and 2002 mean and the 2003 anomaly as predicted by LPJmL, MOD17+, ANN, FPA, and FPA+LC. 117

Alphabetic list of frequently used acronyms

ANN: Artificial Neural Network upscaling of GPP, TER, and NEP based on flux tower measurements, FAPAR from MODIS, and meteorological data (Papale and Valentini 2003, Vetter et al. 2007)

APAR: Absorbed Photosynthetic Active Radiation (=FAPAR x PAR) ; [MJ/yr]

Biome-BGC: terrestrial ecosystem model that also models nitrogen dynamics (Thornton 1998)

DBF: Deciduous Broadleaf Forest

EBF: Evergreen Needleleaf Forest

ECMWF: European Centre for Medium-Range Weather Forecasts (refers to the meteorological reanalysis product)

ENF: Evergreen Needleleaf Forest

FAPAR: Fraction of Absorbed Photosynthetic Active Radiation

fAPAR: see FAPAR

FPA: FAPAR based Productivity Assessment; empirical GPP model based on the JRC-FAPAR product and GPP data from CarboEurope sites (Jung et al., submitted, see Chapter 5)

FPA+LC: FAPAR based Productivity Assessment + Land Cover; empirical GPP model based on the JRC-FAPAR product and GPP data from CarboEurope sites with separate functions for different vegetation types (Jung et al., submitted, see Chapter 5)

fPAR: see FAPAR

GLC2000: Global Land Cover 2000; global 1km land cover product (Bartholome and Belward 2005)

GLCC: Global Land Cover Characterisation Database (Loveland et al. 2000)

GPP: Gross Primary Production; [$gC/m^2/yr$]

IWA: Index of Water Availability, defined as the ratio of actual to potential evapotranspiration

JRC: Joint Research Centre of the European Union

LPJ: Lund-Potsdam-Jena global biosphere model (Sitch et al. 2003)

MAT: Mean Annual Temperature; [°C]

MOD17+: extended version of the MOD17 radiation use efficiency model (Running et al. 2004) that simulates GPP, TER, and NEP based on FAPAR from MODIS, meteorological data, and land cover; optimised with CarboEurope flux tower data (Reichstein, 2004).

MODIS: Moderate Resolution Imaging Spectroradiometer; satellite sensor on board of TERRA

NCEP: National Center for Environmental Prediction, refers to the meteorological reanalysis product

NEE: Net Ecosystem Exchange (= -NEP); [$gC/m^2/yr$]

NEP: Net Ecosystem Production (= GPP – TER = NPP – Rh); [$gC/m^2/yr$]

NPP: Net Primary Production (= GPP – Ra); [$gC/m^2/yr$]

ORCHIDEE: 'ORganizing Carbon and Hydrology In Dynamic Ecosystems' (French biosphere model, Krinner et al. 2005)

PAR: Photosynthetic Active Radiation; [MJ/yr]

RUE: Radiation Use Efficiency (=GPP/APAR); [gC/MJ]

Reco: see TER

REMO: Regional (climate) Model, refers to the meteorological data from Feser et al. 2001 where REMO was driven with NCEP reanalysis at the boundaries of the European domain

RMSE: Root Mean Square Error of Prediction

SEAWiFS: Sea-viewing Wide Field-of-view Sensor; satellite sensor on board of SeaStar

SYNMAP: synergetic land cover data set produced for terrestrial carbon cycle studies (Jung et al. 2006, see Chapter 2)

TBM: Terrestrial Biosphere Model

TEM: Terrestrial Ecosystem Model

TER: Terrestrial Ecosystem Respiration

VPD: Vapour Pressure Deficit ; [Pa]

Abstract

Gross primary production (GPP) is the flux of carbon into ecosystems via photosynthesis. GPP constitutes the single largest flux of the carbon cycle and is an important determinant of the net carbon balance. This thesis investigates uncertainties of modelling GPP for Europe. The objectives of the four major chapters are: (1) to construct a global 1km land cover map with improved characteristics for carbon cycle modelling to reduce land cover uncertainties, (2) to identify the relative importance of input data and model structure uncertainties regarding the magnitude, spatial pattern, and interannual variability of simulated GPP, (3) to assess the performance of GPP simulations for forest ecosystems across Europe using eddy covariance based GPP data, and (4) to construct a GPP model by linking remotely sensed vegetation properties with eddy covariance based GPP data and to provide a realistic bound of European GPP by comparison with other data-driven models.

On the continental scale, land cover uncertainties are found to be negligible in comparison to meteorological input data and in particular different model structures (LPJ, Orchidee, Biome-BGC). Three main factors seem to drive discrepant GPP simulations: (1) the representation of crops, (2) the representation of nitrogen dynamics, and (3) the coupling of photosynthesis and canopy conductance and the associated feedback through soil hydrology. Very little agreement of simulated interannual variability among models is highlighted. Interactions of biogeochemical cycles (water-carbon-nitrogen) play possibly a more important role than anticipated but are yet poorly understood.

Three process-oriented models LPJ, Orchidee, and Biome-BGC reproduce qualitatively observed changes of forest GPP along the gradient of mean annual temperature from boreal to Mediterranean climate. The relative root mean square error of prediction is for all three models in the order of 30% but systematic biases of all three models are observed along the climatic gradient. The models underestimate the increase of GPP from boreal to temperate climate, primarily because changes of light absorption (leaf area index) are not adequately modelled, which is likely a consequence of missing nitrogen limitation in LPJ and Orchidee.

The construction of an accurate empirical GPP model is facilitated by regressing the accumulated remotely sensed FAPAR of the growing season period with annual sums of GPP from eddy covariance flux measurements. The new GPP estimate has the advantage of being independent from uncertainties related to meteorological input data, and is compared with a neural network

upscaling method (ANN), a radiation use efficiency model (MOD17+), and LPJ. Consensus regarding the mean annual spatial GPP pattern emerges between the FAPAR based GPP model and ANN (R^2=0.74). Limited agreement exists for the spatial 2003 GPP anomaly pattern also among the three diagnostic models. Mean annual GPP of Europe compares within 5% difference among the three diagnostic models and LPJ if it is accounted for bias from meteorological forcing. Conclusions are drawn regarding the use of data driven models to evaluate process-oriented models.

1 Introduction

1.1 Motivation and Objectives

Projections of the behaviour of the biosphere, in particular the magnitude of climate-carbon cycle feedback are diverging (Denman et al., 2007; Friedlingstein et al., 2006a). Improving the predictability of the evolution of the Earth system is needed to effectively employ mitigation and adaptation strategies to climate change. A better understanding of ecosystem functioning and consequently the terrestrial carbon cycle is currently an intensive field of research.

This thesis was conducted in the frame of the CarboEurope Integrated Project, which aims to understand and quantify the carbon balance of Europe, including the constituent fluxes, the processes shaping the carbon budget, and the uncertainties involved. On the level of continental integration different modelling approaches and observational data streams are brought together to evaluate the current understanding of the carbon cycle of Europe. The top-down approach uses measurements of atmospheric CO_2 in conjunction with an atmospheric transport model in an inversion set-up to estimate spatial and temporal patterns of land surface net carbon exchange. The bottom-up approach comprises terrestrial ecosystem models that aim to mimic a mechanistic functioning of ecosystems. These process-oriented models simulate the entire carbon budget based on atmospheric CO_2, meteorological forcing fields, land cover and soil properties input. Complementary to the process-oriented models, data-oriented modelling approaches are forced with remotely sensed ecosystem properties and tuned using carbon flux measurements from CarboEurope flux tower sites. Ultimatively, model fusion and data – model integration within a carbon cycle data assimilation system (CCDAS) shall provide spatially and temporarily consistent, accurate carbon flux estimates (carbon cycle 'reanalysis'). In principle, CCDAS performs model parameter optimisation and thus corrections of the simulated system trajectory using the observations. CCDAS is an exciting challenge with huge intellectual demand to the community. The success of CCDAS relies on sound (1) quality and quantity of observables that can be assimilated including a good understanding of their uncertainties, (2) the dynamic core, i.e. mechanistic process understanding, (3) numerical schemes of coupling between submodels and regarding the optimisation of model parameters (Raupach et al., 2005; Rayner et al., 2005).

Identifying the major sources of uncertainties of carbon cycle simulations, and evaluating models by linking observational data and different modelling approaches, contributes to reaching the goal of a sound CCDAS.

Central to confidence in CCDAS as well as prognostic ecosystem models that are implemented in Earth system models are realistic simulations of many processes. The net carbon balance is the subtle difference of constituent fluxes:

NBP = GPP – Ra – Rh – H

NBP: net biome productivity (= net carbon balance)
GPP: gross primary production (or gross carbon uptake)
Ra: autotrophic respiration
Rh: heterotrophic respiration
H: carbon loss due to harvest or disturbance (e.g. fire)

GPP is the amount of carbon that is assimilated via photosynthesis. It constitutes the flux of carbon into the ecosystem and is thus a first order constraint of the carbon budget. Effectively reducing uncertainties of the simulated net carbon balance needs systematic and rigorous evaluation of the formulation of major processes. Starting this endeavour with GPP would be logical and systematic.

This thesis deals primarily with uncertainties of modelling GPP for Europe. Uncertainties of model simulations arise from uncertainties in (1) input data, (2) model parameters, and (3) model structure. Parameter uncertainty is currently receiving large attention but is not explicitly investigated here (e.g Knorr and Kattge, 2005; Zaehle et al., 2005), also because parameter uncertainty is formally assessed and minimized within CCDAS. Four major questions are guiding the research presented in this thesis: (1) What are the major sources of uncertainties of process-oriented modelling of GPP for Europe?, (2) How realistic are GPP simulations of process oriented models?, (3) What is the GPP of Europe?, (4) How can uncertainties be reduced effectively? To providing some answers to these questions this PhD thesis cuts across and links

the disciplines of carbon cycle modelling, remote sensing, and ecosystem level measurements of carbon fluxes from eddy co-variance towers.

1.2 Outline of the thesis

Chapter 2 suggests a solution to a common problem of land cover parameterisation of terrestrial ecosystem models. Various global 1km land cover maps from remote sensing are now available but intercomparison reveals large differences among them. In addition the classification legend is not very suitable since certain map classes cannot be easily translated into categories used by the models. Differences among alternative data sets and issues of their classification legend constitute uncertainties for carbon cycle modelling. In chapter 2, an algorithm is presented that allows fusing different global land cover products into a new classification scheme optimised for carbon cycle modelling and thereby exploiting synergies of different land cover mapping approaches. The resulting global 1km land cover product with improved charcteristics for the carbon community, SYNMAP, is being used within the CarboEurope model-intercomparison project (Vetter et al., 2007), and in subsequent chapters of this thesis (Chapter 3 and 5).

Chapter 3 aims to identify major uncertainties of GPP simulations for Europe resulting from input data and model structure. A model simulation experiment is designed that allows the systematic investigation of how alternative land cover data sets, spatial resolution of land cover, meteorological forcing fields, and model structures impact on magnitude, mean spatial pattern, and interannual variations of GPP. The analysis is based on simulations from three process-oriented models: Biome-BGC (Thornton, 1998), LPJmL (Bondeau et al., 2007; Sitch et al., 2003), and Orchidee (Krinner et al., 2005). In comparison to common analysis where generally only the effect of one factor on simulation results is investigated with little or no emphasis on changing spatial or temporal patterns, the adopted approach allows comparing the relative importance of different factors in different dimensions (spatial, temporal, magnitude).

Chapter 4 assesses the capacity of the three process-oriented models Biome-BGC, LPJ, and Orchidee to reproduce observed changes of GPP of forest ecosystems across Europe. It presents the first continental scale data-model comparison study for GPP. The models are confronted with eddy covariance based estimates of GPP and leaf area index (LAI) along a mean annual

temperature gradient running from boreal to Mediterranean climate. A method is proposed that allows identification to what extent erroneous simulations of leaf area and thus light absorption cause biased GPP in the models. However, generalisations of the findings of this study to the European domain is not possible since simulations for the agricultural sector, which covers ~40% of the surface, were not assessed due to a lack of data. Therefore, data-oriented modelling approaches are exploited in the subsequent chapter.

Chapter 5 deals with diagnostic assessment of GPP of Europe. A new approach is introduced that allows estimating GPP based on a remotely sensed biophysical vegetation product (fraction of absorbed photosynthetic active radiation, FAPAR) with the major advantage of being independent from uncertainties that arise from meteorological input data. The results for the European domain are compared with simulations from two independent data-oriented modelling approaches (neural network upscaling, and a radiation use efficiency model) and one process-based model. The synthesis of data-oriented estimates of GPP in conjunction with knowledge gained in chapter 3 on the effect of meteorological input and chapter 4 on model performance for forests, allows the identification of the realistic pattern and magnitude of mean GPP of the European domain. In addition, the analysis allows drawing some conclusions to what extent results from data-driven models can be used to evaluate simulations of process-oriented models.

The main findings and conclusions of this thesis are synthesised in chapter 6.

2 Exploiting synergies of global land cover products for carbon cycle modelling[1]

Abstract

Within the past decade several global land cover data sets derived from satellite observations have become available to the scientific community. They offer valuable information on the current state of the Earth's land surface. However, considerable disagreements among them and classification legends not primarily suited for specific applications such as carbon cycle model parameterizations pose significant challenges and uncertainties in the use of such datasets.

This paper addresses the user community of global land cover products. We first review and compare several global land cover products, i.e. the Global Land Cover Characterisation Database (GLCC), Global Land Cover 2000 (GLC2000) and the MODIS land cover product and highlight individual strengths and weaknesses of mapping approaches. Our overall objective is to present a straight forward method that merges existing products into a desired classification legend. This process follows the idea of convergence of evidence and generates a 'best-estimate' data set using fuzzy agreement. We apply our method to develop a new joint 1 km global land cover product (SYNMAP) with improved characteristics for land cover parameterization of the carbon cycle models that reduces land cover uncertainties in carbon budget calculations.

The overall advantage of the SYNMAP legend is that all classes are properly defined in terms of plant functional type mixtures, which can be remotely sensed and include the definitions of leaf type and longevity for each class with a tree component. SYNMAP is currently used for parameterization in a European model intercomparison initiative of three global vegetation models: BIOME-BGC, LPJ, and ORCHIDEE.

Corroboration of SYNMAP against GLCC, GLC2000 and MODIS land cover products reveals improved agreement of SYNMAP with all other land cover products and therefore indicates the successful exploration of synergies between the different products. However, given that we cannot provide extensive validation using reference data we are unable to prove that SYNMAP is actually more accurate. SYNMAP is available on request from Martin Jung.

[1] Published as: Jung, M., Henkel, K., Herold, M., Churkina, G. (2006): Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sensing of Environment*, 101, 534-553.

2.1 Introduction

Assessing and monitoring the state of the Earth surface is a key requirement for global change research. A suite of global land cover maps have been produced from the remote sensing community (Friedl et al., 2002; Hansen et al., 2000; JRC, 2003; Loveland et al., 2000) and are readily available for a variety of applications. The use of such satellite derived land data sets in modelling studies constitutes a major advance in Earth system science either for improving spatially explicit model parameterization or for model evaluation. However, intercomparisons of land cover products (Giri et al., 2005; Hansen and Reed, 2000; Latifovic and Olthof, 2004 show significant disagreements among them and reveal that the products contain uncertainties. At present, potential users have little guidance which dataset to use and why. Such problems have been recognized and are currently being addressed especially by initiatives like GOFC-GOLD (Global Observation of Forest and Land Cover Dynamics). Driven by international conventions and implementation activities (GCOS, 2004; GEOSS, 2005), GOFC-GOLD in conjunction with the Food and Agricultural Organizations (FAO) and the Global Terrestrial Observing Systems (GTOS) have fostered land cover harmonization and strategies for interoperability and synergy between existing and future land mapping products (Herold et al., in press). See and Fritz, in press proposed to generate an improved hybrid land cover map by fusion of GLC2000 and the MODIS product by taking individual strengths and weaknesses carefully into account. The release of the ENVISAT based GLOBCOVER data set will further enhance the availability of accurate and precise land cover data sets.

Although the land cover community is working hard to supply more data sets with an increasing accuracy, their products are not optimized for direct use in dynamic vegetation and biogeochemical models. The vegetation modellers face a general problem: the classes of the land cover product cannot always be translated into what the models need without introducing uncertainties. Some of the land classes in the classification legends have ambiguous definitions and have to be adjusted before these classes can be parameterized in the models. For example the essential properties of the land cover classes necessary for vegetation model parameterization include degree of woodiness, leaf type, canopy seasonality, and photosynthetic path (C3 or C4). Except different photosynthetic pathways of grasses, these properties are usually definable from remotely sensed data. The land cover legends of existing land cover products, however, have

classes that are not easily translated into this scheme. Essential information is missing; e.g. the IGBP-DISCover class 'Woody Savanna' states that it has 30-60 % trees while it is neither defined which leaf type and phenology is present nor what other land cover class is subdominant. The information about the leaf type and seasonality in particular is crucial for the vegetation model. Since ecophysiological parameters driving the exchange of mass and energy are associated with land cover class in the model, it is vital to minimize uncertainty related to land cover parameterization.

In this paper we present a synergetic land cover product (SYNMAP) with improved characteristics for land cover parameterization of the terrestrial carbon cycle models. SYNMAP provides a relatively simple solution to both problems: disagreements and unsuitable classification legends of existing datasets.

2.2 Aims and objectives

The first goal of our study is to emphasize individual advantages and limitations of available land cover products and show that none of them is perfectly suited for carbon cycle model parameterization. A review the individual land cover mapping approaches will highlight their major strengths and shortcomings in section 2.3. Next, we compare the different land cover products in section 2.4 using agreement maps and indicate considerable disagreement between the data sets that cannot be explained as an artefact of the different legends or acquisition periods. The review and comparison of land cover products shows that none of them is much better than another and serves as justification for producing a land cover product to improve the signal-to-noise ratio by exploring synergies of different land cover mapping approaches. Section 2.5 introduces our method that produces a best-estimate map with a user defined legend based on land cover products from AVHRR, MODIS, and VEGETATION satellite sensors using a fuzzy logic approach. The legend we choose is currently optimized for the biogeochemistry process model BIOME-BGC (Churkina et al., 2003; Running and Hunt, 1993; Thornton, 1998) and adapted to the dynamic vegetation models LPJ (Sitch et al., 2003) and ORCHIDEE, (Krinner et al., 2005). Our new dataset SYNMAP is evaluated in section 2.6 by corroboration with existing land cover products in conjunction with their published validation results. Section 2.7 discusses remaining limitations of our data fusion method and emphasizes advantages of our derived

SYNMAP land cover for carbon cycle modelling. In section 2.8 we propose briefly what the land cover community can do to better satisfy users of global land cover products. Section 2.9 summarizes the main findings of this paper.

2.3 Overview of existing global land cover data sets

2.3.1 General principle

Mapping land cover on global scales is a complex challenge and profited from recent developments in computer science, digital image processing and satellite technology. So far, high-resolution global land cover data sets are available from three optical satellite sensors: NOAA-AVHRR (Hansen et al., 2000; Loveland et al., 2000), TERRA-MODIS (Friedl et al., 2002) and SPOT-VEGETATION (JRC, 2003). The general approach of global land cover mapping is to produce temporal, usually monthly composites from daily or weekly mosaics to minimize cloud cover and data noise due to e.g. atmospheric or viewing angle distortions. Core information originates from multitemporal spectral reflectance measurements and especially vegetation indices (Normalized Difference Vegetation Index, NDVI; Enhanced Vegetation Index for MODIS, EVI) that capture the cycle of plant productivity throughout the year. Monthly composites are then used in conjunction with ancillary data to produce land cover categories according to a defined classification scheme on a regional, e.g. continental window basis or for the whole globe. Major differences between the above mentioned achievements exist that are related to:

(1) Sensor capabilities, i.e. spatial and spectral properties and resolution, repetition rate, and recording of information for data correction and calibration,
(2) Raw data processing, i.e. algorithms for image compositing including cloud detection, directional reflectance calibration, corrections for atmospheric distortions, viewing angle and geographic position,
(3) Acquisition year(s),
(4) Classification system (land cover legend),
(5) Selection of input data for classification,
(6) Classification procedure,

(7) Validation of the final product.

The next sections will give a very brief overview and evaluation of the development of the different land cover products that are available at the global scale with a spatial resolution of 1x1 kilometre and used in this paper (Table 2-1). Please consult the references for detailed information, and the review of (Cihlar, 2000) for general issues of large scale land cover mapping. Validation issues are separately considered in section 2.3.7.

Product	Version	Satellite & Sensor	Acquisition Period	Download
GLCC	2.0	NOAA-AVHRR	April 1992 – March 1993	http://edcdaac.usgs.gov/glcc/glcc.asp
GLC2000	1.0	SPOT-VGT	Nov 1999 – Dec 2000	http://www-gvm.jrc.it/glc2000/
MODIS	V004	TERRA-MODIS	Jan 2001 – Dec 2001	http://duckwater.bu.edu/lc/mod12q1.html

Table 2-1: Global land cover products with 1km spatial resolution used in this study.

2.3.2 NOAA-AVHRR (GLCC)

The development of the Global Land Cover Characterization Data Base (GLCC) pioneered global land cover mapping motivated by the International Geosphere-Biosphere Program (IGBP) in 1992. Global 10-day 1 km resolution AVHRR composites for the period April 1992-March 1993 were recomposited to monthly NDVI data sets (Loveland et al., 2000). Due to the navigation properties of the satellite the geometric accuracy is only in the order of ~3 km. Masks for non-vegetated areas (Barren, Snow and Ice) were produced using thresholds for the maximum NDVI values; water and urban classes were not mapped by the satellite but overlaid from the Digital Chart of the World (DCW, Danko, 1992). Unsupervised clustering of the multitemporal NDVI data was used to separate vegetated areas in 961 land cover regions globally reflecting properties of similar plant productivity and phenological behaviour. All 961 clusters were assigned manually to one of the 94 classes of the Olsen's Global Ecosystem Legend by local experts using a suite of ancillary data such as land use, elevation, and ecoregion maps or high-resolution satellite images. Where necessary, individual clusters were split by overlay analysis with additional data, on-screen digitizing or spectral reclustering. To ensure objectivity several

interpreters worked within the same area. The final map on Olsen's Global Ecosystems legend was reclassified into six additional classification schemes including IGBP-DISCover and Anderson USGS Land Use/Land cover to meet individual needs of intended applications.

2.3.3 SPOT-VEGETATION (GLC 2000)

The Global Land Cover 2000 (GLC 2000) project is a European initiative, coordinated by the Joint Research Institute (JRC) in Ispra, Italy. The project strategy was to define 19 spatial windows for the globe, while for each window a separate group of regional experts were responsible for the mapping. The individual groups were unrestricted for how they produced their product except that the VEGA2000 data set had to be used and the classification scheme had to follow the Land Cover Classification System (LCCS) developed by the FAO (Bartholomé and Belward, in press; Fritz et al., 2003). LCCS is a hierarchical classification structure (Di Gregorio and Jansen, 2000) that allows straightforward and flexible class definitions and aggregation from the regional to the global legend with 22 classes. The VEGA2000 data set consists of daily 1 km SPOT 4 - VEGETATION data (blue, red, near infrared and short wave infrared bands, NDVI) from November 1999 to December 2000. They have been radiometricaly, atmospherically and geometrically corrected while no standardisation of bidirectional reflectance had been applied. Because of a lack of extensive training data, unsupervised clustering in conjunction with various ancillary data was widely used for classification purposes (Bartalev et al., 2003; Eva et al., 2004; Mayaux et al., 2004). Regional land cover mapping strategies and the mosaicing to the global product are described in Fritz et al., 2003. Detailed information is available at the GLC 2000 web page (http://www-gvm.jrc.it/glc2000/publications.htm).

2.3.4 TERRA-MODIS

The MODIS land cover product is based on monthly composites from MODIS Level 2 and 3 data between January and December 2001 and include EVI and spectral bands 1-7; spatial texture, land surface temperature, snow cover and elevation will be used additionally in upcoming versions (Friedl et al., 2002; Strahler et al., 1999). The categorization algorithm (MLCCA, MODIS land cover classification algorithm) is based on supervised artificial neural network

classification in conjunction with decision tree classifiers. A global network of training sites (STEP database, Muchoney et al., 1999) is used to train 10 decision trees until maximal accuracies are attained. All ten decision trees are then used as experts to vote on classes based on the input data so that a classification probability can be estimated that a pixel belongs to each class. Additionally to the voting process prior knowledge is included in the classification procedure that specifies how likely a class is to appear at a geographic location based on ancillary land cover maps, and statistics of class distribution of training sites and overall global distribution of classes. Class label assignment combined classification and prior probabilities to posterior probabilities while the class with highest probability wins. Prior knowledge becomes only decisive if the spectral signature is ambiguous (Friedl et al., 2002). The MODIS land cover product is available with different legends including IGBP-DISCover and is intended to be updated annually.

2.3.5 Advantages and shortcomings of land cover mapping approaches

Individual advantages and limitations of the different approaches relate to the points listed in section 2.3.1 and are briefly evaluated here based on the consultation of the literature. In terms of the quality and amount of used satellite data, we recognize a progression from GLCC to GLC2000 and MODIS land cover product. GLCC is based on poorly or uncorrected raw data, using only monthly NDVI composites that also have some geometric problems. The VEGA dataset of GLC2000 with daily composites of calibrated spectral bands and NDVI offers significantly improved data and more flexibility for classification. A further advantage of the VEGA2000 is the effective geometric correction procedures (Bartholomé and Belward, in press). The input datasets of the MODIS product supersede GLCC and GLC2000 in terms of the spectral properties of the MODIS instrument, specifically designed for land surface mapping. Also, the MODIS data are based on higher spatial resolution of the raw data (250 m / 500 m) and comprehensive strategies of data correction and calibration using additional data collected by the instrument as well as including more spectral bands and additional information (Strahler et al., 1999).

Regarding the applied classification methods MLCCA clearly seems the most sophisticated algorithm. In contrast to GLCC and GLC2000, it is purely objective, reproducible, and operational for the whole globe; thus seems most suitable for change detection. The main limitation of MLCCA is its sensitivity to the training data. Friedl et al., 2002 note: "classification results produced from MODIS data are heavily dependent on the integrity and representation of global land cover in the site data, and substantial ongoing efforts are devoted to maintaining and augmenting the STEP database" (p.300). However, given the large variances and sometimes ambiguous signatures of land cover classes in a global context manual cluster assignment and manipulation by experts as done for GLCC and GLC2000 may produce a better map product. The bottom-up approach of GLC2000, i.e. the individual choice of data pre-processing, usage of ancillary data (e.g. other satellite data), classification method and regional classification legend by project participants, allows accounting for region specific characteristics and landscape complexity at the expense of internal consistency of the final global product. Therefore, the quality of the regional products of GLC2000 also varies according to the quality of the regional experts and so do the areas where the different areas were merged. However, for regional studies the individual tiles of GLC2000 seem to offer the most elaborate representation of land coverage.

In relation to the classification legends and classification systems of land cover products two aspects are important: availability of the products with different legends to meet needs of diverse applications and consistency of the classification system itself. GLCC offers the most flexibility for users in terms of available reclassifications including the Olsen classification with 94 classes. MODIS is also available in different legends, which is not the case for GLC2000. LCCS of GLC2000 is the most advanced and flexible classification system with a clear rationale and standardized definition of the classes. But none of legends of all three global land cover products are easily translated into the land cover classes of vegetation models without introducing uncertainty due to poor definition of mixed classes or a lack of information about leaf type and phenology. For example, the BIOME-BGC model distinguishes between seven vegetation classes: evergreen needleleaf trees, evergreen broadleaf trees, deciduous needleleaf trees, deciduous broadleaf trees, shrubs, C3 and C4 grasses. In terms of carbon cycle modelling an accurate representation of tree coverage and its leaf characteristics is required given that the trees largely determine the carbon budget of an area. In this respect it is not sensible to regard a pixel as forest if it is covered by only 15 % trees as defined in LCCS.

In conclusion, the main advantage of GLCC is its availability with many different legends and high thematic resolution, while the relatively poor input data used for classification constitute its major shortcoming. Although GLC2000 benefits strongly from the use of LCCS and its regional bottom-up approach, its global map lacks some internal consistency associated with the individual mapping initiatives by different project participants. The advantage of the MODIS product is its base on a large amount of high-quality earth observation data and an advanced and operational classification algorithm but, in contrast to the regionally tuned GLC2000 approach, heavily relies on the quality of a comprehensive global training database. However, while it is worth evaluating mapping approaches a map product should be judged according to how good it actually is. Therefore the next section deals with validation efforts of land cover products.

2.3.6 Validation of global land cover products

Determining the accuracy of land cover maps is essential but a poses a challenge to be performed at global scales. Four approaches are used to quantitatively estimate the accuracy of land cover classifications: confidence values of the classifier, comparison with other maps, cross validation with training datasets, and statistically robust spatial sampling and acquisition ground reference information. The latter is regarded as most reliable and will be discussed briefly in the next section. A thorough review of accuracy issues of land cover maps is available in Foody, 2002 and this author emphasizes that: “Despite the apparent objectivity of quantitative metrics of accuracy, it is important that accuracy statements be interpreted with care” (p.196).

Validation reference data for global datasets usually originate from the interpretation of high resolution satellite images (e.g. Landsat and SPOT). The common approach is to calculate error (or contingency) matrices between reference and map data. Three measures are commonly used to describe the map and class specific accuracies: overall, user’s- and producer’s accuracy. Overall accuracy is simply the percentage of correctly classified pixels, commonly calculated as area weighted estimates for the different classes. Class specific accuracies can be reported from two points of view, either from the map or the validation side. The producer’s accuracy of a class is defined as the percentage of validation sites classified correctly, while the user’s accuracy of a class relates to the percentage of map area classified correctly. Both neglect either omission or

commission errors, i.e. whether pixels that should belong to that class are classified as another class or misclassified as that class respectively. Problematic here is that the area of the individual land cover class affects the class specific accuracies (MODIS land cover team (2003)). The classification scheme and class distance (which depends on application) should be considered but usually is not (Mayaux et al., in press). For the purpose of this study, confusion between 'Trees' and 'Barren' is more crucial than between 'Woody savannah' and 'Savannah' but it counts the same in classical confusion statistics. Accuracy statements are therefore sensitive to the level of class aggregation and maps with a high thematic resolution are more likely to be less accurate according to confusion calculations.

The validation initiatives for GLCC (Scepan, 1999), GLC2000 (Mayaux et al., in press), and MODIS (MODIS land cover team (2003)) land cover products have reported overall area weighted accuracies of 67 %, 69 % and 71 % respectively. However, since different databases and approaches were used it must be emphasized that reported accuracy measures are not comparable and should not be regarded as truly robust quantitative estimates. While the GLCC and GLC2000 validation used design-based sample schemes, the MODIS product accuracy assessment is based on a cross-validation – i.e. using several subsets of the training data (which have not been used for the training process) as reference information. It is therefore not possible, to judge which product is better than another in overall or class specific performance.

Given the daunting and expensive task of global land cover product validation, the working group on calibration and validation of CEOS (Committee Earth Observation Satellites) have recently prepared a 'best practise document' to provide thorough validation standards for global land cover datasets CEOS, in press. In this respect, any consistent and operational validation algorithm has to consider the standardized acquisition of reference information to allow for comparative assessment of the validity, strengths and weaknesses of individual datasets (Herold et al., in press).

In the next section we show to what extent GLCC, GLC2000 and MODIS agree or disagree to each other by presenting agreement maps. We will show that there is significant disagreement between the products, which is related to their different land cover mapping procedures rather then due to different legends or periods of satellite data acquisition.

2.4 (Dis)Agreement of GLCC, GLC2000 and MODIS land cover products

2.4.1 Method

To facilitate overlaying of the different maps the data sets need to be co-registered and homogenized ("cross-walked") to a common legend. As base projection we chose simple geographic (latitude/longitude; Plate Carrée) projection with a spatial resolution of 30x30'' (0.008333°) since all data sets were available with this projection. GLC2000 has a slightly deviating spatial resolution of 0.008929° and therefore had to be resampled to 0.008333° using nearest neighbour. We checked the georeference information of the products and found them to be precise - additional co-registration would not improve the spatial match. We clipped all maps to a subset with 43200 columns and 17500 rows to exclude Antarctica, which is not covered by GLC2000. Resampling and subsetting data was done in ENVI 4.0; the remaining image processing and data modelling outlined below was coded in IDL 6.0.

We have defined a simplified legend with nine major classes that accommodate all land cover categories on an aggregated level according to the occurrence of major life forms (SIMPLE-legend). A legend translation table for the original legends is given in Table 2-2. Lumping classes such as combining all forest and savanna types is necessary to account for the diverse classification schemes since e.g. GLC2000 defines forest with > 15 % tree cover while IGBP-DISCover distinguishes between savannas (10-30 %), woody savannas (30-60 %) and forest (> 60%). Equally, LCCS splits shrublands (> 15 % shrub coverage) according to leaf longevity into deciduous and evergreen; in contrast, IGBP-DISCover separates between open (10-60 % coverage) and closed (> 60 %) shrubland. This lumping of classes increases the agreement between the data sets at the expense of thematic precision.

The reclassified data sets of GLCC, GLC2000 and MODIS are then overlaid to produce a map of agreement, revealing where all three, two or none of the maps show equal representation of the land surface. To assess how much of the discrepancy between the maps may be related to land cover change during the acquisition periods of 1992-1993 and 2000/2001, the case that MODIS agrees with GLC2000 but both disagree with GLCC is treated separately and named 'potential

land cover change' with a strong emphasis on 'potential'. This exercise aims to identify whether land cover change is an important factor in explaining discrepancy between maps. To get further indication if land cover change is responsible for the discrepancy between the land cover products we produce an agreement map of the GLCC-IGBP and MODIS-IGBP data sets. This is independent from potential artefacts of the reclassification procedure including its artificial increase of accuracies. The 'potential land cover change' case from the comparison on the SIMPLE-legend basis is used as a mask and overlaid, which allows an approximation of the impact of varying acquisition periods on the discrepancy between GLCC-IGBP and MODIS-IGBP.

SIMPLE	IGBP-DISCover	LCCS
Trees	- Evergreen Needleleaf Forest - Evergreen Broadleaf Forest - Deciduous Needleleaf Forest - Deciduous Broadleaf Forest - Mixed Forests - Woody Savannas - Savannas	- Tree Cover, broadleaved, evergreen - Tree Cover, broadleaved, deciduous, closed - Tree Cover, broadleaved, deciduous, open - Tree Cover, needle-leaved, evergreen - Tree Cover, needle-leaved, deciduous - Tree Cover, mixed leaf type - Mosaic: Tree cover / Other natural vegetation - Tree Cover, burnt
Shrubs	- Closed Shrublands - Open Shrublands	- Shrub Cover, closed-open, evergreen - Shrub Cover, closed-open, deciduous
Grasses	- Grasslands	- Herbaceous Cover, closed-open
Wetlands	- Permanent Wetlands	- Tree Cover, regularly flooded, fresh water - Tree Cover, regularly flooded, saline water - Regularly flooded Shrub and/or - Herbaceous Cover
Barren	- Barren or Sparsely Vegetated	- Sparse Herbaceous or sparse Shrub Cover - Bare Areas
Snow	- Snow and Ice	- Snow and Ice
Crops	- Croplands	- Cultivated and managed areas
Crops/Natural Vegetation Mosaic	- Cropland/Natural Vegetation Mosaic	- Mosaic: Cropland / Tree Cover / Other natural vegetation - Mosaic: Cropland / Shrub or Grass Cover
Urban	- Urban and Built-Up	- Artificial surfaces and associated areas

Table 2-2: Table with translation between the SIMPLE-legend and the IGBP-DIScover (for GLCC and MODIS) and LCCS (GLC2000) legends.

2.4.2 Result

The agreement of the GLCC, GLC2000 and MODIS land cover maps, reclassified to the SIMPLE legend, is presented in Figure 1-1a. All three maps equal each other to 41 %, mainly in areas with extensive tree coverage (e.g. tropical and boreal forest zones), barren (e.g. Sahara and Gobi desert), cropland (e.g. central Europe, India) and snow/ice coverage (Greenland). Further 45 % is related to the agreement of only two maps, while the contribution of the 'potential land cover change' case to that number is 12 %. Still 14 % remain where all three land cover maps disagree. Areas where all maps disagree or only two maps agree seem to be associated with mainly transitional ecozones with mixtures of the three main components trees, shrubs and grasses such as tropical savannas including the Sahel, Mediterranean Europe and tundra.

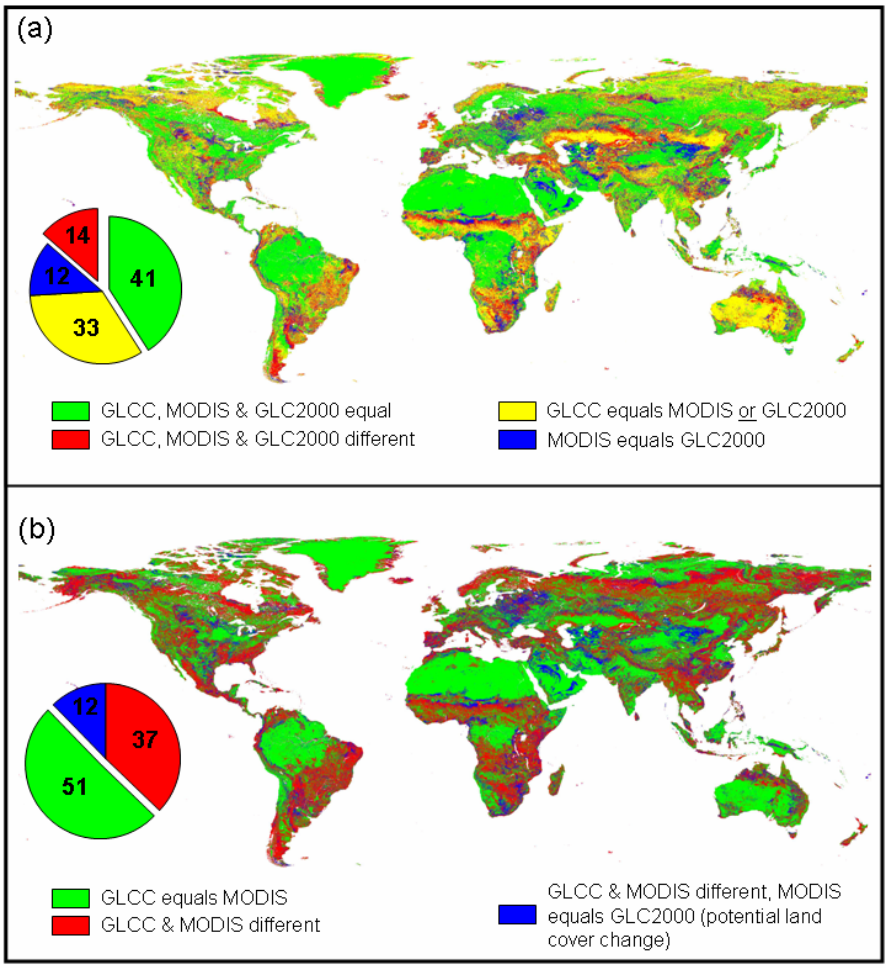


Figure 1-1: Maps of agreement and disagreement between land cover products. (a) GLCC, GLC2000 and MODIS converted to SIMPLE legend. (b) GLCC and MODIS on IGBP-DISCover legend. The pie charts and numbers therein give percentages of the individual cases. Please note that these numbers are not area estimates since the analysis is based on Plate Carrée projection.

The direct comparison of the GLCC and MODIS products on the IGBP DisCover legend in Figure 1-1b reveals that both maps agree for 51 % of land pixels only. Their disagreement cannot be explained by land cover change between the acquisition periods 1993 and 2001 since the 'potential land cover change' case from Figure 1-1a overlaid on the IGBP DisCover agreement map contributes only 12 %.

2.4.3 Discussion

While large scale homogeneous areas seem to be represented reliably in all land cover products, large discrepancies are apparent in heterogeneous landscapes. There are several reasons for that. Firstly, mapping a continuum of e.g. trees, shrubs and grasses into discrete categories is problematic (Foody, 2002). For instance, if one map shows grassland and another shrubland they disagree although both may be right. This 'mixed unit' problem seems a major challenge for all coarse scale land cover mapping efforts because the heterogeneity of the landscape structure is more detailed than the resolution of the satellite sensor (Smith et al., 2002). When several land cover types are present within a pixel, the signature becomes ambiguous and the classification very sensitive to the method. The sensors with fine spatial resolution (e.g. IKONOS or even LANDSAT) are capable to resolve the landscape structure, but are still less effective for mapping large land areas. Perhaps, it is hoped that consistent global land cover information of Landsat-type data may be developed and made available. For map intercomparison, geographic misregistration also becomes crucial in heterogeneous terrain (Foody, 2002). Another general problem is the low separability between certain classes such as e.g. grass- and shrublands, whose statistical signatures overlap in the multidimensional space; and if both land cover types are present it becomes even more challenging.

From this simple comparison of land cover products, we draw two conclusions. First, there is significant discrepancy between them that cannot be explained by different classification schemes or acquisition dates. We therefore disagree with Giri et al., 2005 who relate the main disagreements between GLC2000 and MODIS land cover products to different classification schemes. Land cover change between 1993 and 2000 cannot explain the discrepancies between GLCC and GLC2000 or MODIS. There is common agreement within the land cover mapping community that the accuracies of the individual land cover products are insufficient to detect land

cover change reliably. The deviation of the three products seems largely to be caused by the methodology and input data. Second, a land cover map derived by blending the original products would be a significant improvement, given that much information and confidence is gained in areas where at least two maps agree. Land cover change between the acquisition periods is not an issue as demonstrated, especially not because two third of the land cover information would originate from 2000/2001, with GLCC from 1993 as an additional contributor of important information.

2.5 Land cover data fusion – exploring synergies between land cover products

2.5.1 General principle

Since the aim is to combine several land cover classifications to a best estimate land cover map for a new user defined legend, a flexible method is needed capable of handling differing classification schemes and their fuzziness. For each product, land cover has been classified into a limited number of classes while the boundaries between thematically adjacent classes are, to some extent, arbitrary drawn and a question of definition and accuracy, which vary between land cover products. Also, the separability of adjacent or mixed classes is very limited due to overlapping signatures so that class assignment becomes very sensitive to the classification algorithm. The basic idea of fuzzy logic here seems a welcome rationale by blurring the boundaries of land cover classes (Ahlqvist et al., 2003). In principle, our method consists of two steps: the definition of the desired classification legend and secondly, to link the defined legend classes with the legend classes of the original maps by assigning affinity scores between them. This provides a score for each pixel and defined class, while the class with the highest score wins that can in principle be understood as a voting process of the input data sets. The next sections describe how this principle has been implemented to produce a land cover map with an optimized legend for terrestrial carbon cycling modelling – SYNMAP.

2.5.2 Definition of the target legend

There are two important requirements for the definition of the target legend. First, it is dependent on the application of the land cover map, in our case land cover parameterization of the global vegetation model where it is important to have leaf type and longevity of trees specified. Second, the information for all desired classes must be available from the input data sets.

We have chosen the target legend to consist of the vegetation types used by BIOME-BGC, because it is both directly translatable into the ecophysiological parameters crucial for carbon cycle and can be easily classified using remotely sensed data (Running et al., 1995) To account for the co-occurrence of vegetation types we define a class as single or a combination of maximal two vegetation life forms (16 categories). For each land cover class, which has a tree component, leaf type and longevity are to be specified. It results in 48 classes, while 36 of them are associated with tree coverage (see Table 2-3). We assume that the indicated SYNMAP class covers more than 50 % of the pixel; in the case of mixed classes the indicated class combination is maximal relative to all other class possibilities. The legend of SYNMAP is, in contrast to existing products, flexible and ideal for upscaling to a coarser grid cell size with fractional estimates of the vegetation types used by BIOME-BGC as demonstrated in section 2.7.

2.5.3 Selection and pre-processing of input data sets

To allow for most suitable land cover characterization, it is advantageous to use different classification schemes to perform cross-mapping of classes. We chose five different data sets: the USGS and IGBP legend for GLCC and the PFT and IGBP legend for the MODIS product; GLC2000 is only available with the LCCS legend but goes twice in the calculation. We decided not to use the land cover product from the University of Maryland (UMD) based on AVHRR 1992-1993 data (Hansen et al., 2000) because we wanted to keep the majority of information from 2000/2001 and preliminary visual inspection and statistical comparison with other land cover products suggested that only little information can additionally be gained from this data set. The UMD classification has further not been validated and is less widely used in the research community.

Life form	Tree leaf type	Tree leaf longevity	SIMPLE
Trees	Needle	Evergreen	Trees
Trees	Needle	Deciduous	
Trees	Needle	Mixed	
Trees	Broad	Evergreen	
Trees	Broad	Deciduous	
Trees	Broad	Mixed	
Trees	Mixed	Evergreen	
Trees	Mixed	Deciduous	
Trees	Mixed	Mixed	
Trees & Shrubs	Needle	Evergreen	
Trees & Shrubs	Needle	Deciduous	
Trees & Shrubs	Needle	Mixed	
Trees & Shrubs	Broad	Evergreen	
Trees & Shrubs	Broad	Deciduous	
Trees & Shrubs	Broad	Mixed	
Trees & Shrubs	Mixed	Evergreen	
Trees & Shrubs	Mixed	Deciduous	
Trees & Shrubs	Mixed	Mixed	
Trees & Grasses	Needle	Evergreen	
Trees & Grasses	Needle	Deciduous	
Trees & Grasses	Needle	Mixed	
Trees & Grasses	Broad	Evergreen	
Trees & Grasses	Broad	Deciduous	
Trees & Grasses	Broad	Mixed	
Trees & Grasses	Mixed	Evergreen	
Trees & Grasses	Mixed	Deciduous	
Trees & Grasses	Mixed	Mixed	
Trees & Crops	Needle	Evergreen	Crops/Natural Vegetation Mosaic
Trees & Crops	Needle	Deciduous	
Trees & Crops	Needle	Mixed	
Trees & Crops	Broad	Evergreen	
Trees & Crops	Broad	Deciduous	
Trees & Crops	Broad	Mixed	
Trees & Crops	Mixed	Evergreen	
Trees & Crops	Mixed	Deciduous	
Trees & Crops	Mixed	Mixed	
Shrubs & Crops	-	-	
Grasses & Crops			
Crops	-	-	Crops
Shrubs	-	-	Shrubs
Shrubs & Grasses	-	-	
Shrubs & Barren	-	-	
Grasses	-	-	Grasses
Grasses & Barren	-	-	
Barren	-	-	Barren
Urban & Built-Up	-	-	Urban
Permanent Snow & Ice	-	-	Snow

Table 2-3: SYNMAP legend defined by dominant life form assemblage and tree leaf attributes. The last column gives a translation to the SIMPLE legend.

IGBP-DISCover and LCCS have 17 and 23 classes respectively. The USGS classification scheme has 24 classes with several 'mixed' classes; while the PFT legend consists of only 11 classes, defined clearly by the dominant life form and leaf attributes for forest classes, which contributes important information in cases where several possibilities of desired 'mixed classes' are possible. For example if the GLCC and MODIS maps on IGBP legend indicate a 'Savanna' type land cover, the PFT map may specify whether it is a 'Trees & Shrubs' or 'Trees and Grasses' mosaic by indicating 'Shrubland' or 'Grassland' respectively.

In addition to the land cover products we use AVHRR-CFTC (Continuous Fields of Tree Cover) data sets to contribute information on leaf type and phenology for tree classes. The CFTC products give fractional information of leaf type and leaf longevity for each pixel with tree coverage with two layers for both attributes respectively: needleleaf/broadleaf and evergreen/deciduous, while each pair sums up to percentage tree cover. They have been derived from monthly AVHRR NDVI composites of the 1992-1993 period (see section 2.3) using spectral unmixing (DeFries et al., 1999). Using CFTC data in addition to the land cover classification products has the advantage that information on leaf type and longevity can be estimated for areas with tree coverage that are below the forest threshold ('savanna' or 'woodland' classes) and hence lack important information on leaf characteristics of present trees. Therefore, AVHRR CFTC were converted into a leaf type and leaf longevity map by dividing the data into three discrete classes, i.e. needleleaf, broadleaf, mixed and evergreen, deciduous, mixed respectively. The data were rescaled to 100 % so that e.g. percentage of needleleaf plus percentage of broadleaf equal 100 % and not percentage of tree cover as in the original data layers. The three classes are equally spaced, hence mixed is assigned if neither needleleaf nor broadleaf or evergreen nor deciduous exceed 66 %.
For overlaying all data sets were prepared at 30'' spatial resolution in Plate Carrée projection as outlined in section 2.3 and the final product (SYNMAP) has equally a resolution of 30''.

2.5.4 Definition of affinity scores

Affinity scores link our defined legend classes with the legend classes of the original products and therefore approximate the thematic distance of the classes. Affinity scores are defined for life form, leaf type and leaf longevity separately. Each class of each original land cover data set is

assigned to one or more 'target' classes with a score between zero and four according to semantic rules (Table 2-4). A land cover class X gets assigned a zero to target class Y if both are independent from each other such as 'Barren' and 'Trees'. It contributes two points if land cover class X is a component of class Y such as 'Grassland' and 'Shrubs & Grasses', or 'Evergreen needleleaf forest' and 'Mixed leaf type'. Four points are given if class X matches class Y, for instance 'Evergreen needleleaf forest' and 'Trees' or 'Open shrubland' and 'Shrubs & Grasses'. One or three points are given in some cases when flexibility is needed and indicate minor or major components of class X in class Y. The definition of scores requires knowledge of the original classification schemes and is to some degree subjective. All tables with affinity scores are presented in appendix 1.

Land cover class example	Semantic rule	Affinity score	Target class example
Woody savanna	'is not'	0	'Barren'
	'has minor parts of'	1	'Grasses'
	'has parts of'	2	'Trees'
	'has major parts of'	3	'Trees & Grasses'
	'is'	4	'Trees & Shrubs'

Table 2-4: Definition of affinity scores according to semantic rules. The example uses the IGBP-Discover class 'Woody savanna'.

2.5.5 Calculation of SYNMAP

The calculation is done in two steps: the first step is related to dominant life forms, the second step performs estimation of leaf attributes if a tree component is present in the life form assemblage. Applying the score tables for all five data sets a total score is calculated for each target land cover class for each pixel. To be consistent that each product contributes the same amount of information, the GLC2000 data set was used twice. In addition to GLCC, GLC2000 and MODIS land cover products, information on leaf type and leaf longevity is added from the reclassified CFTC maps. A pixel gets assigned the class for which the total score is maximal. The concept is illustrated in Figure 2-2 in conjunction with Tables 2-5 and 2-6.

Instead of calculating the total scores for each target class along the pixel vector across the different land cover products (6 addends, 7 addends for leaf attributes), we place a 3x3 window around each pixel. Hence, information for each target class is accumulating from 54 addends

(3x3x6) (63 for leaf attributes), while the center of the pixel is weighted by eight to make it equally important to the sum of all the remaining window pixels. Including neighbouring pixels in the calculation has several advantages. It accounts to some extent for potential inaccuracies of the georeferencing of the original products and artefacts from resampling of the original sensor data from which the land cover maps were produced. Secondly, it improves the reliability of the resulting map in very heterogeneous areas such as the Mediterranean (see section 2.4). It further reduces the chance that two or more classes receive the same maximal score and therefore acts to force a decision.

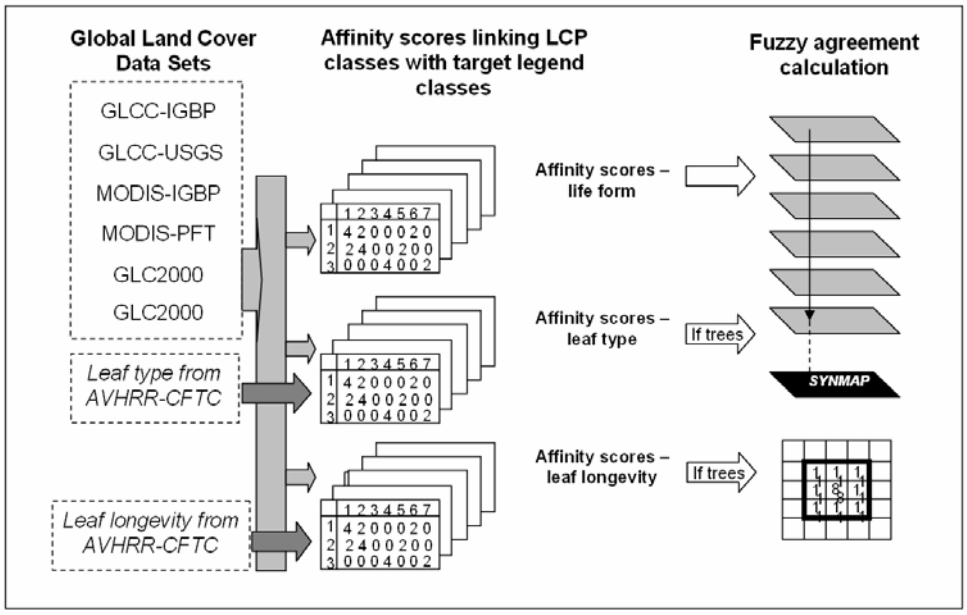


Figure 2-2: Principle of the data fusion method. The legends of land cover products are linked with the target legend using affinity scores for life forms, leaf type and leaf longevity. Leaf type and leaf longevity maps from CFTC data contribute additionally to the calculation of leaf attributes. Leaf attributes are calculated if trees are present in the life form assemblage. Each land cover product contributes the same amount of information; for MODIS and GLCC two different reclassifications are used while GLC2000 is only available with one global legend and therefore counts double in the calculation. Fuzzy agreement of the different maps is calculated for a 3x3 pixel window with the center pixel being weighted by eight according to Equation 2-1.

The choice of the SYNMAP class is therefore made according to the following equation that calculates the total score for each life form (T) of the SYNMAP legend for grid cell with coordinates i and j of SYNMAP:

$$S_{Total}^{T}(i,j) = 8\sum_{M=1}^{6} S_M^T(i,j) + \sum_{M=1}^{6} S_M^T(i-1,j) + \sum_{M=1}^{6} S_M^T(i+1,j) + \sum_{M=1}^{6} S_M^T(i,j-1) + \sum_{M=1}^{6} S_M^T(i,j+1) +$$
$$+ \sum_{M=1}^{6} S_M^T(i-1,j-1) + \sum_{M=1}^{6} S_M^T(i+1,j+1) + \sum_{M=1}^{6} S_M^T(i-1,j+1) + \sum_{M=1}^{6} S_M^T(i+1,j-1)$$

Equation (2-1): Calculation of total score for each life form of the SYNMAP legend.

$S_{Total}^{T}(i,j)$ - total score for life form T of SYNMAP legend;

$S_M^T(i,j)$ - affinity score for the life form in the grid cell (i,j) of the existing land cover map M assigned to class T of SYNMAP legend (Appendix 1);

T - life form of SYNMAP legend (Table 2-3);

M - existing land cover maps (GLCC-USGS, GLCC-IGBP, MODIS-PFT, MODIS-IGBP, 2xGLC2000)

i - current row of pixels

j - current column of pixels

The life form with the maximum total score $S_{Total}^{T}(i,j)$ is chosen as the best estimate of the life form in grid cell (i,j) of SYNMAP. In case 'trees' are present in the life form assemblage leaf type and leaf longevity are estimated according to the same principle but integrating over seven land cover maps because CFTC data are included.

Data Set	Original land cover class	Trees & Shrubs	Trees & Grasses	Shrubs	Grasses	Shrubs & Grasses
GLCC-IGBP	'Open Shrubland'	1	0	2	2	4
GLCC-USGS	'Mixed Shrubland/Grassland'	1	1	2	2	4
MODIS-IGBP	'Woody Savanna'	4	3	1	1	1
MODIS-PFT	'Shrub'	2	0	4	0	2
GLC2000	'Herbaceous Cover'	0	2	0	4	2
GLC2000	'Herbaceous Cover'	0	2	0	4	2
Total Score		8	8	9	13	**15**

Table 2-5: Calculation example for the best estimate of life form assemblage along the pixel vector of the land cover data sets. GLC2000 is taken twice in the calculation to be consistent that each product supplies the same amount of information. The class indicated by each layer contributes scores to the target legend classes. The target legend class with the highest score wins, here 'Shrubs & Grasses'.

Data Set	Original land cover class	Needle	Broad	Mixed
GLCC-IGBP	'Mixed Forest'	2	2	4
GLCC-USGS	'Mixed Forest'	2	2	4
MODIS-IGBP	'Evergreen Needleleaf Forest'	4	0	2
MODIS-PFT	'Evergreen Needleleaf Forest'	4	0	2
GLC2000	Tree Cover, broadleaf, deciduous	0	4	2
GLC2000	Tree Cover, broadleaf, deciduous	0	4	2
CFTC	Broadleaf	0	4	2
Total Score		12	16	**18**

Table 2-6: Calculation example for leaf type. Note that CFTC data contribute information in addition to the land cover products. The calculation for leaf longevity (not shown) operates the same way.

In case two or more life form classes receive the same maximal score, the decision which life form class wins is made by a priority rule; in our case it is simply the ascending order of class values. If leaf type or leaf longevity cannot be assigned because more than one leaf class has the same maximal score a decision matrix defines the winning leaf attributes (Table 2-7). If no information for leaf attributes is available (i.e. maximal score is zero), both leaf type and leaf longevity are set to 'mixed'. This compromise introduces uncertainty, which is fortunately small since this case is very rare and applies only to vegetation mosaics with some tree coverage so that only part of the leaf attribute information of that class is biased.

The next section presents the SYNMAP data set that we have derived from our fuzzy logic based method and evaluates the success of the data fusion process.

Leaf longevity/ leaf type	'Mixed'	'Deciduous/ Broadleaf'	'Evergreen/ Needleleaf'
'Mixed'	-	'Deciduous'	'Evergreen'
'Deciduous/ Broadleaf'	'Broadleaf'	-	'Mixed'
'Evergreen/ Needleleaf'	'Needleleaf'	'Mixed'	-

Table 2-7: Decision matrix for leaf type (below diagonal) and longevity (above diagonal) in case two leaf classes receive the same score.

2.6 SYNMAP evaluation

On a qualitative basis, SYNMAP agrees with what we would expect, reproducing the present day ecological zones on Earth (Figure 2-3). The indication of mixed leaf types of the tree and grass savanna at the fringe of the Sahel to the Sahara is, however, erroneous and should be deciduous broadleaf (raingreen trees). Here, the life form assemblage 'Trees & Grasses' with no information about leaf attributes was indicated by the input data sets, so that leaf type and longevity were set to mixed to minimize the error. For model parameterization, the difference between 'Deciduous broadleaf trees and grasses' and 'Evergreen needleleaf and deciduous broadleaf trees and grasses' is rather small and acceptable. We assess the success of the data fusion method and hence the reliability of SYNMAP quantitatively through intermap comparison with its input data GLCC, GLC2000 and MODIS land cover data sets and link the results to the validation efforts of the original data sets.

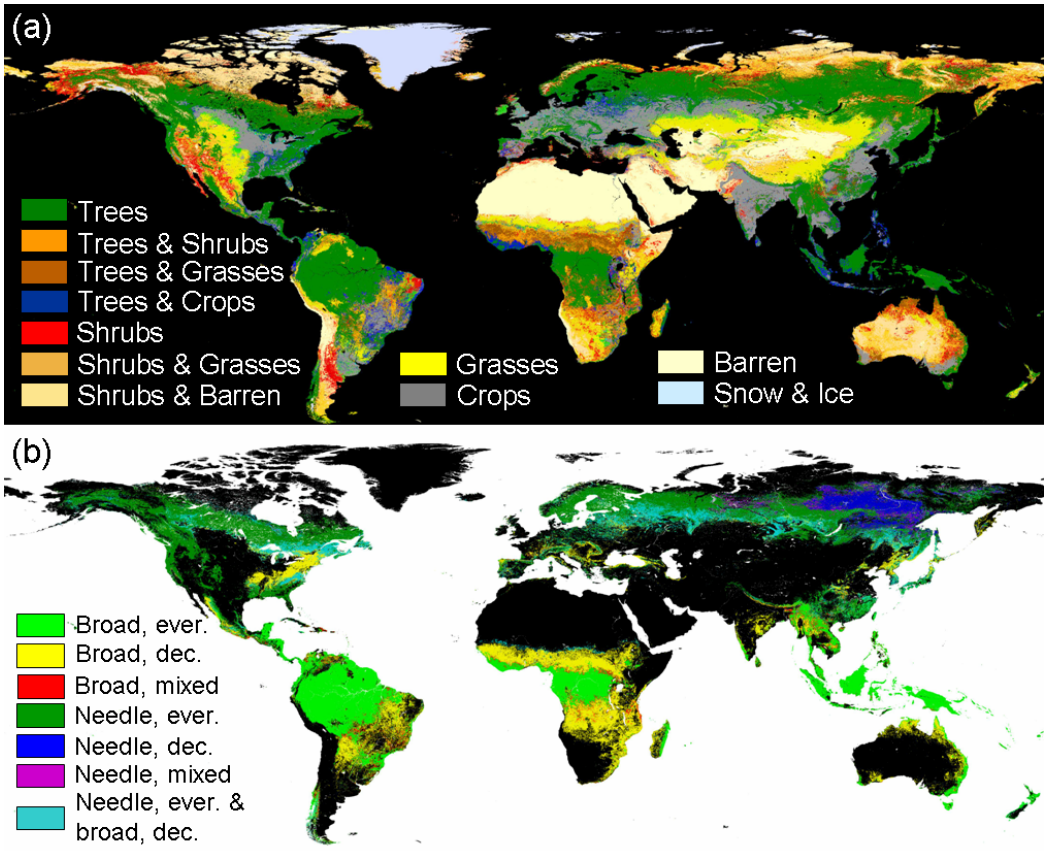


Figure 2-3: The SYNMAP data set. (a) Life form assemblages. 'Shrubs & Crops', 'Grasses & Barren' and 'Urban' have too little extent and are invisible on that scale. (b) Leaf attributes of trees. All areas are displayed where a tree component is present. 'Needle leafed, mixed longevity', 'Mixed leaf types, evergreen' and 'Mixed leaf types deciduous' have too little extent and are invisible on that scale.

To make GLCC, GLC2000, MODIS and SYNMAP land cover products comparable in terms of classification schemes we reclassified the data sets into the SIMPLE legend (see Tables 2-2 and 2-3). Minor bias due to different original classification legends for some classes remains after this step since e.g. the SYNMAP class 'Shrubs & Barren' is considered as 'open shrubland' and hence included in 'Shrubs' while it also overlaps with the class 'Sparsely vegetated' which is included in 'Barren'. We then calculate pixel-based confusion matrices between each combination pair; water and wetland areas have been excluded here. From confusion matrices we derive overall consistency for life forms and leaf attributes separately, which we define as the percentage of pixels where both maps agree on the class. To get an estimate of the mean overall consistency of a map we simply average the calculated consistency estimates where the considered map was a comparison partner (Equation 2-2). The consistency for the tree leaf attributes (Evergreen needleleaf, Evergreen Broadleaf, Deciduous Needleleaf, Deciduous Broadleaf and Mixed) is related to confusion within forest classes only (where both maps indicate forest) that have information on leaf type and longevity since forests were defined according to different tree cover threshold by IGBP-DIScover (>60 %) and LCCS (>15 %). We use the term 'consistency' between land cover products to avoid 'accuracy' which would not be entirely correct given that we do not provide validation against reference data. The consistency measure is calculated analogous to accuracy from confusion tables as in common remote sensing practise.

Mean $C_a = (C_{ab}+C_{ac}+C_{ad})/3$

Equation (2-2): Calculation of mean consistency for map a.

Indices a – d are the maps (GLCC, GLC2000, MODIS, SYNMAP)

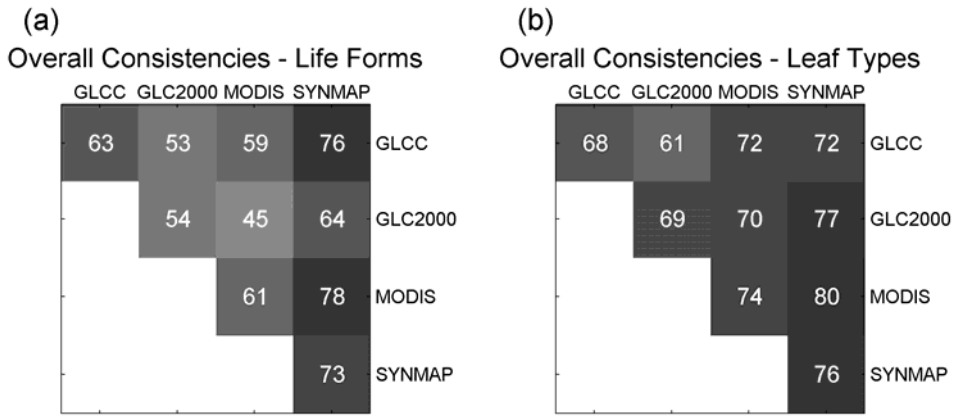


Figure 2-4: Overall consistency between GLCC, GLC2000, MODIS and SYNMAP based on major life forms (SIMPLE legend) (a) and leaf attributes (b). Average overall consistencies of the maps are given along the diagonal.

Figure 2-4 presents the calculated consistency between data sets for major life forms and leaf attributes. Among all map pairs, SYNMAP agrees best with each original land cover product regarding the occurrences of major life forms and leaf attributes. Average map specific consistencies are presented along the diagonal.

We further calculate class specific consistencies for each land cover product based on the SIMPLE legend, which we define here as the percentage of 'right' classified pixels of a class (where both maps agree) relative to the number of pixels that have been classified as that class by either of the maps (Equation (2-3)). Therefore, the class consistency includes both omission and commission errors and will be much lower than usually reported User's or Producer's accuracies. Average class specific consistencies are calculated according to Equation (2-2).

$$CC_{ab} = \frac{n_{ab}}{n_a + n_b - n_{ab}} \times 100$$

Equation (2-3): Calculation of class consistency.

CC_{ab} is class consistency of a class between the two maps *a* and *b*

n_{ab} is number of pixels mapped as respective class by both maps (*a* and *b*)

n_a is number of pixels mapped as respective class by map *a*

n_b is number of pixels mapped as respective class by map *b*

Indices a – d are the maps (GLCC, GLC2000, MODIS, SYNMAP)

In addition to the consistencies derived from intermap comparison, we calculate 'reference' accuracies for GLCC, GLC2000 and MODIS products from published confusion matrices in their individual validation analysis papers (Mayaux et al., in press; Scepan, 1999; MODIS land cover team (2003)). We determine these land cover product specific reference accuracies in the same way and using the same SIMPLE reclassification as in the map corroboration analysis. Linking the accuracies from the original validation data of the land cover products to the consistency estimates derived from intermap comparisons gives insights whether it is possible to approximate the class accuracies of the land cover products by map to map corroborations and further provides a benchmark of the reliability of a particular class.

Figure 2-5 shows class specific consistencies derived from intermap comparison and reference accuracies calculated from published validation exercises of GLCC, GLC2000 and MODIS. The map corroborations suggest that SYNMAP has the highest average class consistencies for all classes except 'Deciduous Needleleaf' where MODIS is slightly higher. For 'Trees', 'Crops', 'Grasses' and 'Deciduous Broadleaf', and to a lesser extent 'Shrubs' and 'Crops/Natural Vegetation Mosaic' classes SYNMAP exhibits particular enhanced agreement with the other maps.

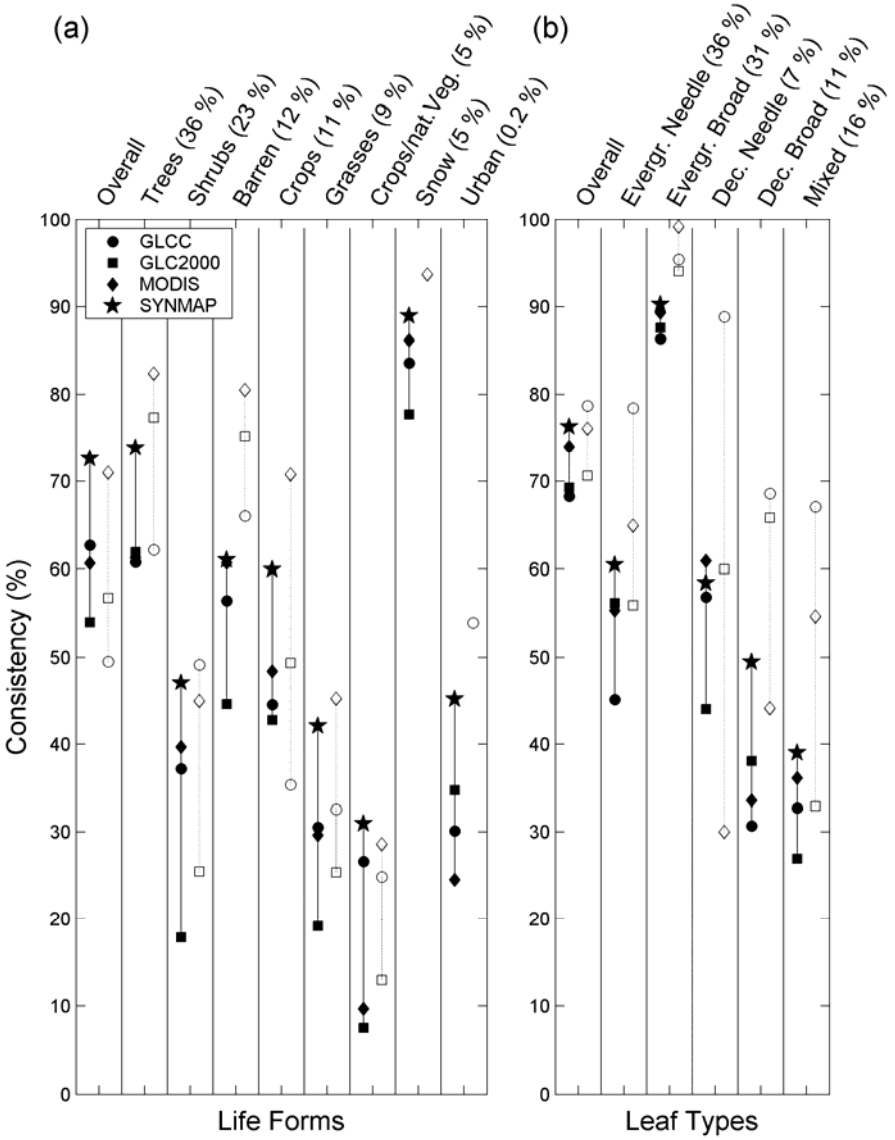


Figure 2-5: Average class specific consistencies for GLCC, GLC2000, MODIS and SYNMAP derived from intermap comparison (filled markers) and 'ground truth' class accuracies derived from published confusion matrices for GLCC, GLC2000 and MODIS based on validation data (not filled markers). The area extent (in lat/lon pixels) within SYNMAP is given as percentages in brackets. 'Snow' and 'Urban' classes were not sampled by all validation exercises. Note that no area weighting has been done to calculate overall accuracies from the published confusion matrices with ground truth. Area weighting would shift overall accuracies up.

For major life forms, the range of class consistencies calculated from map to map comparisons and class accuracies from validation agrees well on a qualitative basis, except for 'Barren', indicating that map corroboration approximates overall and class accuracies. The accuracies for leaf attributes derived from validation data for GLCC, GLC2000 and MODIS appear to be higher than the corresponding consistencies from intermap comparisons. The discrepancy for 'Barren' and leaf attributes may be an artefact of the sensitivity of ground truth sites to their location (see section 2.3.6).

Figure 2-5 also reveals a general pattern of class reliability of current global remote sensing based land cover maps. Areas with trees, snow and barren are accurately recognized. Croplands can be considered to be ok. Shrublands, Grasslands and urban area are problematic; Croplands in association with natural vegetation are very problematic as well. Regarding leaf classes of trees, evergreen broadleaf is very well represented in maps, evergreen and deciduous needleleaf are well reproduced, while deciduous broadleaf and mixed seem more uncertain.

2.7 Limitations and advantages of SYNMAP

A major concern about SYNMAP is that we cannot provide rigorous validation against reference data at the moment. However, we have shown that our data fusion method was successful so that SYNMAP represents the best agreement between GLCC, GLC2000 and the MODIS land cover product that all had been validated individually. We further cannot state strict definitions of the SYNMAP classes involving thresholds such as at what exact percentage of tree coverage a forest class is mapped. This is a critical point, but according to our knowledge and experience this is more of a theoretical issue; in practise the current capabilities of global land cover mapping cannot provide such precise information accurately anyway. One may also see the definition of affinity scores as a further weak point. We assigned the affinity scores between original land cover product and SYNMAP classes ad hoc according to semantic rules and our knowledge; they have not been derived in a purely objective, quantitative way and do not take individual strengths and weaknesses of the products into account.

However, we have provided a simple and useful solution to a common problem with land cover data sets. The proposed data fusion method can be applied to produce a map for a specific

purpose with a desired legend from existing maps. Essential is here, that the legends of the existing products are studied and appropriately linked to the target legend via affinity scores using semantic rules. If one is confident that a certain input data set is of significantly higher quality then a higher weight can be given to that product, e.g. it goes twice into the calculation of class specific scores. However, giving weight to individual classes of input products is very dangerous and should not be applied since higher weighted classes have a higher probability to be mapped in the product and thus will erroneously occur too often. Knowledge about class specific accuracies of the products can, in conjunction with the calculated fuzzy agreement, be used to generate a map of confidence of the final product. Especially, when the target legend contains classes that are different from the legends of the input data, several target legend classes may get an equal score for a pixel, i.e. there is no unambiguous class to be mapped. To tackle this problem it is advantageous to use many data sets with different legends and to include neighbouring pixels with a smaller weight into the calculation to increase the number of estimates and therefore confidence.

Despite the remaining limitations of SYNMAP we are confident that we have produced a data set that is better suitable to parameterize carbon cycle models than existing ones. The SYNMAP legend is well suited and uncertainties resulting from cross-walking the map classes to the model vegetation classes are reduced. Particularly important is that, in contrast to existing products, leaf characteristics are defined for mixed forest and mosaic classes of trees with other vegetation or cropland given that biophysical parameters are associated with these traits in the model. We believe SYNMAP to be more accurate than existing land cover products since it makes use of synergies between different land cover mapping approaches that all have their individual strengths and limitations and a blend of the different maps should enhance the signal-to-noise ratio, which is indeed indicated by intermap comparisons.

While SYNMAP is specifically developed for carbon cycle modelling it may be suitable for other applications. Disregarding the accuracy of a land cover product, its applicability for a specific purpose depends on which classes are considered in respect to the application requirements. SYNMAP resembles the information content of the legends of its input data sets but has more distinct definitions of which vegetation types are present in mixed classes. Thus, SYNMAP can be aggregated to a coarser model grid cell which contains fractions of the globally most relevant

plant functional types without having mixed classes: evergreen needleleaf tree, evergreen broadleaf tree, deciduous needleleaf tree, deciduous broadleaf tree, shrub, grass, crop, and also urban, barren and water. Hence, if such information is sufficient for a particular purpose, SYNMAP may be used for this. If, however, information with greater thematic detail is needed other sources such as the regional tiles of GLC2000 with region specific legends or data sets from higher resolution satellite imagery (e.g. Landsat) should be consulted. Currently, there is no 'wetland' class in SYNMAP because global vegetation models do not deal with wetlands explicitly at present. Furthermore, the 'wetland' class in land cover products from optical remote sensing is rather uncertain since the sensor is sensitive to vegetation coverage rather than wetness.

The SYNMAP data set is currently used in a carbon cycle model intercomparison initiative of the CARBOEUROPE-IP project that also investigates the effect of using different land cover products for parameterizations. Preliminary analysis of the model results show that SYNMAP based calculations for all carbon budget variables plot in between those for MODIS and GLC2000, which gives us with further confidence of the quality of SYNMAP (Jung, unpublished).

2.8 The way forward from a user's perspective

The common problem is that a single product will never be perfectly suited for all applications either in terms of spatial coverage, accuracy and/or in terms of the legend. Therefore it is important that the land cover community generates more and increasingly accurate products. But what would be really desirable and would constitute a great contribution to Earth system modelling is an online tool where users can design their own legends of the land cover product they need. This is very challenging but it may possible for the future. Major steps that need to be taken include:

(1) fostering interoperability of land cover products and developing a common land cover language such as LCCS that links product legends in a quantitative or semi-quantitative way

(2) studying the accuracy and individual strengths and weaknesses of existing products thoroughly and identifying problem areas and classes that need to be remapped

(3) developing a sophisticated data fusion method that makes use of (1) and (2)
(4) compiling an extensive data base of reference data and providing operational validation.
(5) setting up storage and computing facilities, collecting all data and providing a web based interface

As an increasing array of land cover classifications and continuous fields data is accumulating an operational method that merges these data in a desired classification legend and provides a validation against reference data straight away would result in more accurate products that ultimatively best satisfy the users. Especially the 'difficult areas', i.e. the transitional and heterogeneous zones would be better mapped when various sources are used with their individual strengths and weaknesses taken into account and when the user himself can decide in what he is more interested by designing the target legend in a hierarchical way (e.g. according to LCCS). The work of Herold et al., 2006, Fritz and See, 2005 and See and Fritz, in press is promising and already goes in that direction.

2.9 Summary and conclusions

Initiatives of global land cover mapping have used diverse approaches and data from different satellite sensors with varying degrees of raw data corrections and manual manipulation during the classification process. It is not surprising that they produced different results and it is currently not possible to judge which map is more suitable for a specific purpose.

Based on the individual validation efforts and inter-map comparison of GLCC, MODIS and GLC2000 land cover products we identified problematic areas. Trees (woodlands), snow covered as well as bare areas seem reliably mapped, while discrepancies exist within forest classes such as confusion between deciduous broadleaf and mixed forest. Croplands are well represented, as long as they are not grouped with natural vegetation. The 'Cropland/Natural Vegetation Mosaic' class is the least reliable land cover category. In addition, significant uncertainties are associated with grass- and shrubland classes.

Spatially, regions of disagreement between data sets are primarily related to transitional zones with mixed classes; land cover change between acquisition periods is found to be of second order significance. Problem areas and problem classes are connected and mainly related to two issues: class separability and spatial heterogeneity, as well as mapping a continuum transition with a discrete classification scheme. The first issue may be regarded as a general problem of optical remote sensing in discriminating certain categories that have large intra-class variance of their multitemporal spectral signatures that overlap with other categories such as shrub- and grasslands or wetlands. The second issue is one of cartographic standards. When different maps give various estimates for areas with mixed classes, all may be right and wrong to some extent since maps are forced to fit the real world into categories being very sensitive to classification algorithms and representation of mixed cartographic units.

Classification schemes and class definitions are problematic in several ways. More or less arbitrary thresholds are applied to distinguish between classes such as open and closed forest. Mixed classes especially lack clear definitions, partly because it is not possible given the limitations of global land cover mapping. Different initiatives used classification schemes, which make inter-map comparison challenging and only possible on an aggregate class level. Especially problematic for users of land cover products is that classification schemes may be not flexible enough for their application because important information, e.g. the specification of leaf attributes for ‘savanna’ type for carbon cycling modelling parameterization, is missing.

Motivated by the disagreement and classification legends unsuitable for terrestrial carbon cycle modelling of existing products we developed a method that generates a new global land cover map (SYNMAP) dedicated for terrestrial carbon cycle modelling with biogeochemistry and dynamic vegetation models. SYNMAP has been already successfully applied for parameterization of the BIOME-BGC, LPJ and ORCHIDEE models. The data fusion process blends different global land cover products based on fuzzy agreement and allows the definition of a desired target legend.

Pixel based intercomparison of SYNMAP, GLCC, GLC2000 and MODIS land cover products at an aggregated class level reveals highest overall and class specific consistency for SYNMAP and therefore indicates the successful exploration of synergies between products. Although we

believe that our data set is more accurate than existing land cover products, a thorough test of SYNMAP is only possible using a comprehensive database of ground truth information. The current developments for more operational and harmonized land cover observations, both in situ and satellite based, will provide consistent and continuous representations of the Earths land cover: in more spatial detail, with more flexible legends, and with robust and comparable accuracy statements (Herold et al., in press).

The general problem of using land cover data sets in carbon cycle modelling is the subject of ongoing research. Categorical maps such as SYNMAP are straightforward to use in many global models but a limited number of discrete classes is problematic for two main reasons. Firstly, it depends on how land cover is perceived (i.e. definition of classes) and there is likely to be a mismatch between the product and the model assumptions even if the class is called the same. Secondly, actual land cover is a continuum and usually a composition of different vegetation types is present in a grid cell (usually about 0.5 degree for large scale carbon cycle modelling). Vegetation continuous fields products are an attempt to tackle this problem and for some models it is sensible to implement such data sets. Beer, 2005 applied MODIS vegetation continuous fields data for the boreal region in LPJ model and found significant improvements in comparison to discrete modelled natural vegetation coverage. Another alternative is to go beyond a complete physiognomic land cover description and focus on measurable traits and biophysical variables which require other types of carbon cycle models. For instance the Turgor-Mesic-Sclerophyll scheme is a framework that links canopy leaf property with vegetation structure and resource availability (Berry and Roderick, 2002). This scheme describes properties of leaves, not vegetation types or species. It has been already successfully applied for the Australian vegetation. It is highly desirable to conduct more studies on alternative methods to carbon cycle modelling.

5.10 Appendix – Affinity scores for life forms, leaf types and leaf longevities

Affinity scores for life forms

IGBP-DisCover legend (GLCC & MODIS)	Water	Trees	Trees & Shrubs	Trees & Grasses	Trees & Crops	Shrubs	Shrubs & Grasses	Shrubs & Crops	Shrubs & Barren	Grasses	Grasses & Crops	Grasses & Barren	Crops	Barren	Urban	Snow
Water	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Evergreen needleleave forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Evergreen broadleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Deciduous needleleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Deciduous broadleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Mixed forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Closed shrublands	0	0	2	0	0	4	2	2	2	0	0	0	0	0	0	0
Open shrublands	0	0	1	0	0	2	4	2	4	2	1	1	0	2	0	0
Woody savannas	0	2	4	3	2	1	1	0	0	1	0	0	0	0	0	0
Savannas	0	1	3	4	2	2	2	1	1	2	1	0	0	0	0	0
Grasslands	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Permanant wetlands	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Croplands	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Urban & built-up	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
Cropland/natural vegetation mosaic	0	2	2	2	4	2	2	4	1	2	4	1	2	0	0	0
Snow & ice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
Barren or sparsely vegetated	0	0	0	0	0	1	2	0	3	1	0	3	0	4	0	0
PFT legend (MODIS)	Water	Trees	Trees & Shrubs	Trees & Grasses	Trees & Crops	Shrubs	Shrubs & Grasses	Shrubs & Crops	Shrubs & Barren	Grasses	Grasses & Crops	Grasses & Barren	Crops	Barren	Urban	Snow
Water	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Needleleaf evergreen tree	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Broadleaf evergreen tree	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Needleleaf deciduous tree	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Broadleaf deciduous tree	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Shrub	0	0	2	0	0	4	2	2	2	0	0	0	0	0	0	0
Grass	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Cereal crop	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Broadleaf crop	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Urban & built-up	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
Snow & ice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4

Barren or sparsely vegetated	0	0	0	0	0	1	2	0	3	1	0	3	0	4	0	0
USGS- legend (GLCC)	Water	Trees	Trees & Shrubs	Trees & Grasses	Trees & Crops	Shrubs	Shrubs & Grasses	Shrubs & Crops	Shrubs & Barren	Grasses	Grasses & Crops	Grasses & Barren	Crops	Barren	Urban	Snow
Urban & built-up land	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
Dryland cropland & pasture	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Irrigated cropland & pasture	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Mixed Dryland/irrigated cropland & pasture	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Cropland/Grassland Mosaic	0	0	0	0	0	2	1	2	0	2	4	1	2	0	0	0
Cropland/Woodland Mosaic	0	2	1	1	4	0	0	1	0	0	1	0	2	0	0	0
Grassland	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Shrubland	0	0	2	0	0	4	2	2	2	0	0	0	0	0	0	0
Mixed shrubland/grassland	0	0	1	1	0	2	4	1	1	2	1	1	0	0	0	0
Savanna	0	2	4	4	2	2	2	1	1	2	1	1	0	0	0	0
Deciduous broadleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Deciduous needleleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Evergreen broadleaf forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Evergreen needleleave forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Mixed forest	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Water bodies	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Herbaceous Wetland	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Wooded Wetland	0	4	3	2	1	1	1	0	1	1	0	0	0	0	0	0
Barren or sparsely vegetated	0	0	0	0	0	1	2	0	3	1	0	3	0	4	0	0
Herbaceous tundra	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Wooded tundra	0	2	4	3	1	2	2	1	1	2	1	1	0	0	0	0
Mixed tundra	0	2	3	4	1	2	3	1	1	2	1	1	0	0	0	0
Bare ground tundra	0	0	0	0	0	0	0	0	2	0	0	2	0	4	0	0
Snow or ice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
LCCS legend (GLC2000)	Water	Trees	Trees & Shrubs	Trees & Grasses	Trees & Crops	Shrubs	Shrubs & Grasses	Shrubs & Crops	Shrubs & Barren	Grasses	Grasses & Crops	Grasses & Barren	Crops	Barren	Urban	Snow
Tree Cover, broadleaved, evergreen	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Tree Cover,	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0

broadleaved, deciduous, closed																
Tree Cover, broadleaved, deciduous, open	0	4	3	3	2	1	1	0	0	1	0	0	0	0	0	0
Tree Cover, needle-leaved, evergreen	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Tree Cover, needle-leaved, deciduous	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Tree Cover, mixed leaf type	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Tree Cover, regularly flooded, fresh water	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Tree Cover, regularly flooded, saline water	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Mosaic: Tree Cover / Other natural vegetation	0	2	4	4	2	2	2	1	1	2	1	1	0	0	0	0
Tree Cover, burnt	0	4	2	2	2	0	0	0	0	0	0	0	0	0	0	0
Shrub Cover, closed-open, evergreen	0	0	2	0	0	4	3	2	3	0	0	0	0	0	0	0
Shrub Cover, closed-open, deciduous	0	0	2	0	0	4	3	2	3	0	0	0	0	0	0	0
Herbaceous Cover, closed-open	0	0	0	2	0	0	2	0	0	4	2	2	0	0	0	0
Sparse herbaceous or sparse shrub cover	0	0	1	1	0	2	2	1	4	2	1	4	0	3	0	0
Regularly flooded shrub and/or herbaceous cover	0	0	2	2	0	4	4	2	2	4	2	0	0	0	0	0
Cultivated and managed areas	0	0	0	0	2	0	0	2	0	0	2	0	4	0	0	0
Mosaic: Cropland / Tree Cover / Other natural vegetation	0	2	1	1	4	1	1	2	1	1	2	1	2	0	0	0
Mosaic: Cropland / Shrub and/or grass cover	0	0	1	1	1	2	2	4	1	2	4	1	2	0	0	0
Bare Areas	0	0	0	0	0	0	0	0	2	0	0	2	0	4	0	0
Water Bodies	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Snow and Ice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
Artificial surfaces and associated areas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0

Affinity scores for leaf types and leaf longevities

IGBP-DisCover legend (GLCC & MODIS)	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Water	0	0	0	0	0	0
Evergreen needleleaved forest	4	0	2	4	0	2
Evergreen broadleaf forest	0	4	2	4	0	2
Deciduous needleleaf forest	4	0	2	0	4	2
Deciduous broadleaf forest	0	4	2	0	4	2
Mixed forest	2	2	4	2	2	4
Closed shrublands	0	0	0	0	0	0
Open shrublands	0	0	0	0	0	0
Woody savannas	0	0	0	0	0	0
Savannas	0	0	0	0	0	0
Grasslands	0	0	0	0	0	0
Permanant wetlands	0	0	0	0	0	0
Croplands	0	0	0	0	0	0
Urban & built-up	0	0	0	0	0	0
Cropland/natural vegetation mosaic	0	0	0	0	0	0
Snow & ice	0	0	0	0	0	0
Barren or sparsely vegetated	0	0	0	0	0	0
PFT legend (MODIS)	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Water	0	0	0	0	0	0
Needleleaf evergreen tree	4	0	2	4	0	2
Broadleaf evergreen tree	0	4	2	4	0	2
Needleleaf deciduous tree	2	0	2	0	4	2
Broadleaf deciduous tree	0	4	2	0	4	2
Shrub	0	0	0	0	0	0
Grass	0	0	0	0	0	0
Cereal crop	0	0	0	0	0	0
Broadleaf crop	0	0	0	0	0	0
Urban & built-up	0	0	0	0	0	0
Snow & ice	0	0	0	0	0	0
Barren or sparsely vegetated	0	0	0	0	0	0
USGS- legend (GLCC)	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Urban & built-up land	0	0	0	0	0	0

Dryland cropland & pasture	0	0	0	0	0	0
Irrigated cropland & pasture	0	0	0	0	0	0
Mixed Dryland/irrigated cropland & pasture	0	0	0	0	0	0
Cropland/Grassland Mosaic	0	0	0	0	0	0
Cropland/Woodland Mosaic	0	0	0	0	0	0
Grassland	0	0	0	0	0	0
Shrubland	0	0	0	0	0	0
Mixed shrubland/grassland	0	0	0	0	0	0
Savanna	0	0	0	0	0	0
Deciduous broadleaf forest	0	4	2	0	4	2
Deciduous needleleaf forest	4	0	2	0	4	2
Evergreen broadleaf forest	0	4	2	4	0	2
Evergreen needleleave forest	4	0	2	4	0	2
Mixed forest	2	2	4	2	2	4
Water bodies	0	0	0	0	0	0
Herbaceous Wetland	0	0	0	0	0	0
Wooded Wetland	0	0	0	0	0	0
Barren or sparsely vegetated	0	0	0	0	0	0
Herbaceous tundra	0	0	0	0	0	0
Wooded tundra	0	0	0	0	0	0
Mixed tundra	0	0	0	0	0	0
Bare ground tundra	0	0	0	0	0	0
Snow or ice	0	0	0	0	0	0
LCCS legend (GLC0000)	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Tree Cover, broadleaved, evergreen	0	4	2	4	0	2
Tree Cover, broadleaved, deciduous, closed	0	4	2	0	4	2
Tree Cover, broadleaved, deciduous, open	0	4	2	0	4	2
Tree Cover, needle-leaved, evergreen	4	0	2	4	0	2
Tree Cover, needle-leaved, deciduous	4	0	2	0	4	2
Tree Cover, mixed leaf type	2	2	4	2	2	4
Tree Cover, regularly flooded, fresh water	0	0	0	0	0	0
Tree Cover, regularly flooded, saline water	0	0	0	0	0	0
Mosaic: Tree Cover / Other natural vegetation	0	0	0	0	0	0

Tree Cover, burnt	0	0	0	0	0	0
Shrub Cover, closed-open, evergreen	0	0	0	2	0	1
Shrub Cover, closed-open, deciduous	0	0	0	0	2	1
Herbaceous Cover, closed-open	0	0	0	0	0	0
Sparse herbaceous or sparse shrub cover	0	0	0	0	0	0
Regularly flooded shrub and/or herbaceous cover	0	0	0	0	0	0
Cultivated and managed areas	0	0	0	0	0	0
Mosaic: Cropland / Tree Cover / Other natural vegetation	0	0	0	0	0	0
Mosaic: Cropland / Shrub and/or grass cover	0	0	0	0	0	0
Bare Areas	0	0	0	0	0	0
Water Bodies	0	0	0	0	0	0
Snow and Ice	0	0	0	0	0	0
Artificial surfaces and associated areas	0	0	0	0	0	0
Leaf type map from AVHRR CFTC	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Needleleaved (>66%)	4	0	2	0	0	0
Broadleaved (>66%)	0	4	2	0	0	0
Mixed leaf type (33-66%)	2	2	4	0	0	0
Leaf longevity map from AVHRR CFTC	Needleleaved	Broadleaved	Mixed leaf type	Evergreen	Deciduous	Mixed leaf longevity
Evergreen (>66%)	0	0	0	4	0	2
Deciduous (>66%)	0	0	0	0	4	2
Mixed leaf longevity (33-66%)	0	0	0	2	2	4

3 Uncertainties of modelling GPP over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models[2]

Abstract

Continental to global scale modelling of the carbon cycle using process based models is subject to large uncertainties. These uncertainties originate from the model structure and uncertainty in model forcing fields, however, little is known about their relative importance. A thorough understanding and quantification of uncertainties is necessary to correctly interpret carbon cycle simulations, and guide further model developments.

This study elucidates the effects of different state-of-the-art land cover and meteorological data set options, biosphere models on simulations of gross primary productivity (GPP) over Europe. The analysis is based on: (1) three different process oriented terrestrial biosphere models (Biome-BGC, LPJ, Orchidee) driven with the same input data, and one model (Biome-BGC) driven with (2) two different meteorological data sets (ECMWF, REMO), (3) three different land cover data sets (GLC2000, MODIS, SYNMAP), and (4) three different spatial resolutions of the land cover (0.25° fractional, 0.25° dominant, 0.5° dominant). We systematically investigate effects on the magnitude, spatial pattern, and interannual variation of GPP.

While changing the land cover map or the spatial resolution has only little effects on the model outcomes, changing the meteorological drivers and especially the model results in substantial differences. Uncertainties of the meteorological forcings,affect particularly strongly interannual variations of simulated GPP.

By decomposing modeled GPP into their biophysical and ecophysiological components (absorbed photosynthetic active radiation (APAR) and radiation use efficiency (RUE) respectively) we show that differences of interannual GPP variations among models result primarily from differences of simulating RUE. Major discrepancies appear to be related to the feedback through the carbon-nitrogen interactions in one model (Biome-BGC) and water stress

[2] To be published as: Jung, M., Vetter, M., Herold, M., Churkina, G., Reichstein, M., Zaehle, S., Cias,P., Viovy,N., Bondeau, A., Chen, Y., Trusilova, K., Feser, F. and Heimann, M.: Uncertainties of modelling GPP over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models. *Global Biogeochemical Cycles,* in press.

effects, besides the modeling of croplands. We suggest clarifying the role of nitrogen dynamics in future studies and revisiting currently applied concepts of carbon-water cycle interactions regarding the representation of canopy conductance and soil processes.

3.1 Introduction

The terrestrial biosphere constitutes a major part of the global carbon cycle and receives large attention in terms of climate change mitigation due to its carbon sequestration potentials (e.g. Prentice et al., 2000). Within the past decades terrestrial biosphere models (TBMs) have been developed to reproduce and predict carbon stocks and fluxes of the land on continental to global scales (Cramer et al., 2001; McGuire et al., 2001). TBMs require a range of input (or driving) data, most importantly meteorological, soil and land cover information. Current input data are of heterogeneous nature and origin and modellers need to make a choice between alternative driver data sets. The quality of these inputs will have an effect on the accuracy of carbon budget calculations. However, the extent of the effects has not yet been quantified systematically. It is further recognized that uncertainties of TBMs themselves are still rather large, both in terms of parameter-based (e.g. Zaehle et al., 2005), and model structure related uncertainty (e.g. Kramer et al., 2002; Morales et al., 2005, Moorcroft, 2006). To develop robust estimates of the behaviour of the biosphere in the future, a thorough understanding of input data effects and model uncertainties should lead to a critical review of current modelling performances and avenues to improve known limitations.

Changing the model inputs or changing the model itself means changing the results, but the question is: by how much and in which dimension? Previous studies had looked at individual aspects such as how the spatial resolution, the choice of the meteorological data set, or parameter uncertainty influences carbon flux simulations, concentrating primarily on net primary production (NPP) (Hicke, 2005; Kimball et al., 1999; Knorr and Heimann, 2001; Turner et al., 2000; Zaehle et al., 2005; Zhao et al., 2006). The studies differed in the scale from regional to global, and in the way they quantified the effects while generally ignoring effects on spatial and temporal patterns. No systematic study has yet been done that allows to judge how different options in the model set-up affects the magnitude, spatial, and temporal patterns of carbon flux simulations. It is of key importance to elucidate what really matters, i.e. to identify first and second order factors.

Such knowledge subsequently allows us to improve efficiently our abilities towards accurate estimates of the global carbon budget.

In this paper we present a systematic study that shows how the choice of the model inputs (land cover map, spatial land cover resolution, meteorological data set), and the choice of the process oriented carbon cycle model itself affect the magnitude, spatial, and temporal patterns of gross primary productivity (GPP) simulations over Europe. We do not aim to identify which data set or model is best but we discuss how these factors constitute limitations on large scale GPP modelling and how we could improve GPP simulations. GPP is the amount of carbon assimilated by plants via photosynthesis, the process that is believed to be among the best understood within ecosystem carbon cycle modelling. In TBMs, GPP represents the flux how carbon enters the system, and which controls many other processes in the models. If GPP is simulated incorrectly, this error propagates to the other carbon budget variables. GPP is thus a good indicator for the effects of different model-set ups on simulations of the carbon cycle.

3.2 Biosphere models and driver data set options

3.2.1 Terrestrial biosphere models

We use three state of the art terrestrial carbon cycle models: LPJ (Sitch et al., 2003), Orchidee (Krinner et al., 2005), and Biome-BGC (Running and Hunt, 1993; Thornton, 1998).

LPJ is a dynamic global vegetation model (DGVM) and originates from the BIOME model family (Haxeltine et al., 1996; Kaplan et al., 2003; Prentice et al., 1992). It simulates the distribution of plant functional types, and cycling of water and carbon on a quasi-daily time-step. LPJ has been used in numerous studies on responses and feedbacks of the biosphere in the Earth System (e.g. Brovkin et al., 2004; Lucht et al., 2002; Schaphoff et al., 2006; Sitch et al., 2005). The version of LPJ used for these calculations has been adapted to account for a realistic treatment of croplands using a crop functional type approach (Bondeau et al., 2007).

The Orchidee DGVM (Krinner et al. 2005) is used as the land surface scheme of the French earth system model IPSL-CM4. It evolved through the unification of the soil vegetation atmosphere

transfer model SECHIBA (de Rosnay and Polcher, 1998; Ducoudre et al., 1993) and the terrestrial carbon model STOMATE (Viovy et al. 1997 ; Friedlingtein 1998). The biophysical processes (photosynthesis, surface energy budget) simulations operate on a half hourly, the carbon dynamics simulations (allocation, respiration, ageing on a daily time step.

Biome-BGC was designed to study biogeochemical processes and has been applied and tested in various studies (e.g. Churkina and Running, 1998; Churkina et al., 2003; Kimball et al., 2000; Kimball et al., 1997; Vetter et al., 2005). It resulted from the generalisation of a stand model for coniferous forests (Forest-BGC, Running, 1994; Running and Gower, 1991) to other vegetation types. It is the only model considered here that includes a nitrogen cycle. As Orchidee, Biome-BGC treats to date croplands as productive grasslands.

3.2.2 Meteorological and land cover forcings

The requirements of our model intercomparison on meteorological driver data constitute (1) a consistent temporal coverage of several decades, (2) a daily resolution, and (3) an adequately high spatial resolution better than half by half degree. These requirements are met by ERA 40 reanalysis from ECMWF (1961-2000; (ECMWF, 2000) and simulations by the regional climate model REMO (Jacob and Podzun, 1997; Feser et al., 2001). REMO was driven by 6-hourly reanalysis from the National Centers for Environmental Prediction (NCEP, Kalnay et al., 1996; Kistler et al., 2001) from 1948 until 2005 at the boundaries of the European domain. The REMO simulations have a substantially higher spatial resolution (50 by 50km) than the original T62 NCEP data (approximately 2°) and can be regarded as improved NCEP reanalysis. The REMO dataset was chosen to drive all models because it extents until 2005; a prerequisite for a concomitant study on the 2003 heat wave (Vetter et al., 2007).

We chose to use three global 1km remote sensing based land cover products that became recently available: the MODIS product (Friedl et al., 2002), Global Land Cover 2000 (GLC2000, Bartholome and Belward, 2005), and SYNMAP (Jung et al., 2006). SYNMAP has been produced as a synergy of various existing land cover products including GLC2000 and MODIS, and was used to drive all three models since its plant functional type (PFT) based classification legend meets better the requirements of biosphere models.

We test the effect of prescribing land cover with different spatial detail using a fractional representation of different PFTs within a 0.25° grid cell as well as the dominant PFT with 0.25° and 0.5° spatial resolution.

3.3 Experimental design

3.3.1 Modelling strategy

We adopt a straightforward strategy where we define a reference set-up which consists of the following combination: the model Biome-BGC is forced with the REMO meteorology, and SYNMAP land cover with PFT fractions in a 0.25° grid cell. Subsequently, we change one of the components at a time: either the model, or the meteorological data set, or the land cover data set, or the spatial resolution. We then compare the simulations with the modified set-up to the reference one to quantify the effect of the changed component on the magnitude, spatial pattern and temporal variation of GPP.

Figure 3-1 displays the modelling strategy in more detail. We make the following changes from the reference set-up to yield alternative realisations: (1) spatial land cover resolution: 0.25° and 0.5° dominant vegetation type; (2) land cover map: GLC2000 and MODIS; (3) meteorological forcing: ECMWF ERA 40; and the carbon cycle model: LPJ and Orchidee. We do not consider effects due to different soil water holding capacity (WHC) data because of a lack of alternative data sets. Investigating the model's sensitivity to 50% changes of WHC across 12 sites in Europe is the scope of active research. A detailed modelling protocol that contains information on regulations of model spin-up and transient runs as well as other input data which are kept fixed for all runs such as atmospheric CO_2 concentration, soil and elevation data sets is available in Vetter et al., in preparation; Vetter et al., 2007 and from the homepage (http://www.bgc-jena.mpg.de/bgc-systems/projects/ce_i/index.shtml).

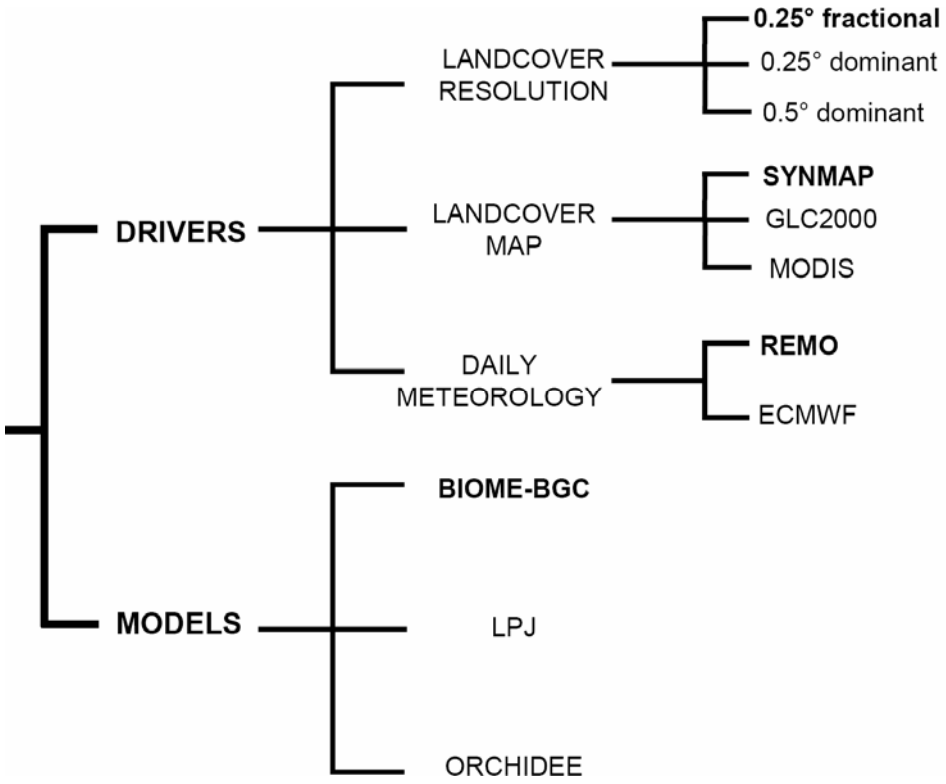


Figure 3-1: Simulation strategy to assess model performance differences due to the choice of the driver data set and carbon cycle model. The ends of the tree to the right present the different options that we consider. The combination of the reference set-up is in bold. From this reference set-up only one component is changed at a time within the branch.

3.3.2 Quantification of effects

All calculations to estimate effects on flux magnitudes, spatial and temporal patterns are based on a 20 year period from 1981 to 2000. We measure the effect on the magnitude in percent as the mean absolute difference of the pixel-based means relative to the mean of the reference (Equation 3-1). To quantify the effect on the spatial pattern we use the variance in percent that is not explained by the squared spatial correlation coefficient between the temporal means of the reference and the alternative realisation (Equation 3-2). We measure the effect on the interannual variability by the variance in percent that is not explained by the squared temporal correlations for each grid cell (Equation 3-3). The mean effect on the temporal patterns is then calculated as the average over all grid cells.

$$EFFECT_{Magnitude} = \frac{\sum_{i=1}^{n} | \overline{AR_i} - \overline{REF_i} |}{\sum_{i=1}^{n} \overline{REF_i}} \times 100 \qquad \text{(Equation 3-1)}$$

$$EFFECT_{Spatial} = 100 - \left(\frac{\sum_{i=1}^{n} (\overline{AR_i} - \overline{\overline{AR}})(\overline{REF_i} - \overline{\overline{REF}})}{\sqrt{\sum_{i=1}^{n} (\overline{AR_i} - \overline{\overline{AR}})^2 \times \sum_{i=1}^{n} (\overline{REF_i} - \overline{\overline{REF}})^2}} \right)^2 \times 100 \qquad \text{(Equation 3-2)}$$

$$EFFECT_{Temporal}\ (i) = 100 - \left(\frac{\sum_{y=1981}^{2000} (AR_y - \overline{AR_i})(REF_y - \overline{REF_i})}{\sqrt{\sum_{y=1981}^{2000} (AR_y - \overline{AR_i})^2 \times \sum_{y=1981}^{2000} (REF_y - \overline{REF_i})^2}} \right)^2 \times 100 \qquad \text{(Equation 3-3)}$$

i: grid cell index

n: number of valid grid cells

y: year

REF: reference modelling set-up

AR: alternative realisation where one component of the reference set-up was changed

The single overbar denotes the grid cell based temporal mean. Two overbars denote the mean over all grid cells of the temporal mean.

3.3.3 Decomposing GPP into absorbed photosynthetic active radiation and radiation use efficiency

In order to gain a better understanding of different GPP simulations by different models we decompose GPP into absorbed photosynthetic active radiation (APAR) and radiation use efficiency (RUE):

GPP = APAR x RUE (Equation 3-4)

The decomposition is carried out for the simulations by different models individually and follows a standard method that has been applied in previous studies (e.g. Bondeau et al., 1999; Ruimy et al., 1999). APAR is calculated from fPAR (fraction of absorbed photosynthetic active radiation) and PAR (photosynthetic active radiation) based on monthly data (Equation (3-5)). fPAR is calculated from modelled LAI according to Lambert-Beer's law assuming a constant light extinction coefficient (k) of 0.5 (Equation 3-6). PAR is assumed to be a constant fraction of 48% of global short wave radiation as simulated by REMO (Equation 3-7). Since we do not account for leaf clumping within the canopy, use constant k and PAR fraction, the derived APAR and RUE values can only be regarded as approximations. However, since we use a consistent methodology the calculated APAR and RUE values are valid for comparison among model simulations.

$$APAR = \sum_{m=1}^{12} fPAR(m) \times PAR(m) \times days(m) \qquad \text{(Equation 3-5)}$$

$$fPAR(m) = 1 - e^{-k \times LAI(m)} \qquad \text{(Equation 3-6)}$$

$$PAR\ (m) = 0.48 \times GRAD\ (m) \qquad \text{(Equation 3-7)}$$

m [-]: month
fPAR [-]: mean fraction of absorbed photosynthetic active radiation
PAR [MJ/m2]: mean photosynthetic active radiation
days: number of days of month m
LAI [m2/m2]: mean (modelled) leaf area index
k [-]: light extinction coefficient (0.5)
GRAD [MJ/m2]: global (short wave) radiation

3.3.4 Investigating the models' response to meteorology

Differences of model behaviour in terms of interannual variability point to different sensitivities to meteorological conditions. Elucidating the sensitivity of simulated GPP to different meteorological variables is difficult since meteorological variables usually covary strongly,

which precludes straightforward separation of the individual effects. We use a principal component analysis (PCA) to effectively reduce the dimensionality of the meteorological input data. We regress the derived variable (first principal component) with simulated variations of GPP to investigate relationship and sensitivity of the models to the meteorology. To better relate the model's response to meteorological conditions we do not use annual data but data from the summer season (June, July, August (JJA)). We first compute mean JJA values for each grid cell and year for temperature, radiation, VPD and precipitation. Subsequently, we remove the variable specific mean and perform a z-score transformation of the data before we compute the PCA in IDL 6.3. The new principal components are then regressed with 'relative' variations of GPP for each grid cell and model. Relative variations are calculated by first subtracting the grid cell based mean and then dividing by the grid cell based mean. We use relative variations because variability generally scales with the flux magnitude, which differs among models. For all grid cells, we calculate Pearson's correlation coefficient, which gives the strength and direction of the relationship between meteorological and GPP variability, as well as the slope of the linear regression line which provides information on the strength of the response, i.e. sensitivity.

3.4 Results and discussion

3.4.1 Order of effects

Table 3-1 summarizes the difference of total GPP of Europe due to alternative model realisations. Changing the meteorological data and the TBM has the largest effects (1.2, 0.9, and 2.1 Gt/yr larger GPP for ECMWF, LPJ, and Orchidee respectively; the reference (Biome-BGC) being 6.2 Gt/yr). The spatial patterns of the difference between reference and alternative realisations are presented in Figure 3-2. The most pronounced effects are again visible for changing the meteorological driver data and the TBM. Major deviations of the ECMWF scenario appear in central, eastern and northern Europe where the ECMWF driven realisation shows substantially higher GPP. The spatial correlation (R^2) of the ECMWF scenario with the reference is 0.67. Changing the model has an even stronger impact on spatial patterns of simulated GPP. In case of Orchidee the correlation (R^2) is 0.54 and for LPJ only 0.2. In the Orchidee simulation the only area where GPP is of similar magnitude is south eastern of the Baltic Sea with otherwise higher GPP. The same area shows decreased GPP in the LPJ simulations. The largest differences to the

LPJ model are found in Western Europe where GPP is up to 1000gC/m^2/yr larger, while small differences are found in north Eastern Europe.

Model set-up	GPP of European domain [GtC/yr]	Difference from reference set-up [GtC/yr]	Difference from reference set-up [%]
Biome-BGC+REMO+SYNMAP+0.25° fractional	6.181	-	-
Biome-BGC+REMO+MODIS+0.25° fractional	6.191	0.010	0.2
Biome-BGC+REMO+GLC2000+0.25° fractional	5.931	-0.250	-4.0
Biome-BGC+REMO+SYNMAP+0.25° dominant	6.551	0.370	6
Biome-BGC+REMO+SYNMAP+0.5° dominant	6.480	0.299	4.8
Biome-BGC+ECMWF+SYNMAP+0.25° fractional	7.397	1.216	19.7
LPJ+REMO+SYNMAP+0.25° fractional	7.031	0.851	13.8
Orchidee+REMO+SYNMAP+0.25° fractional	8.233	2.052	33.2

Table 3-1: Total GPP of European domain as simulated by different model set-ups (1981-2000 mean).

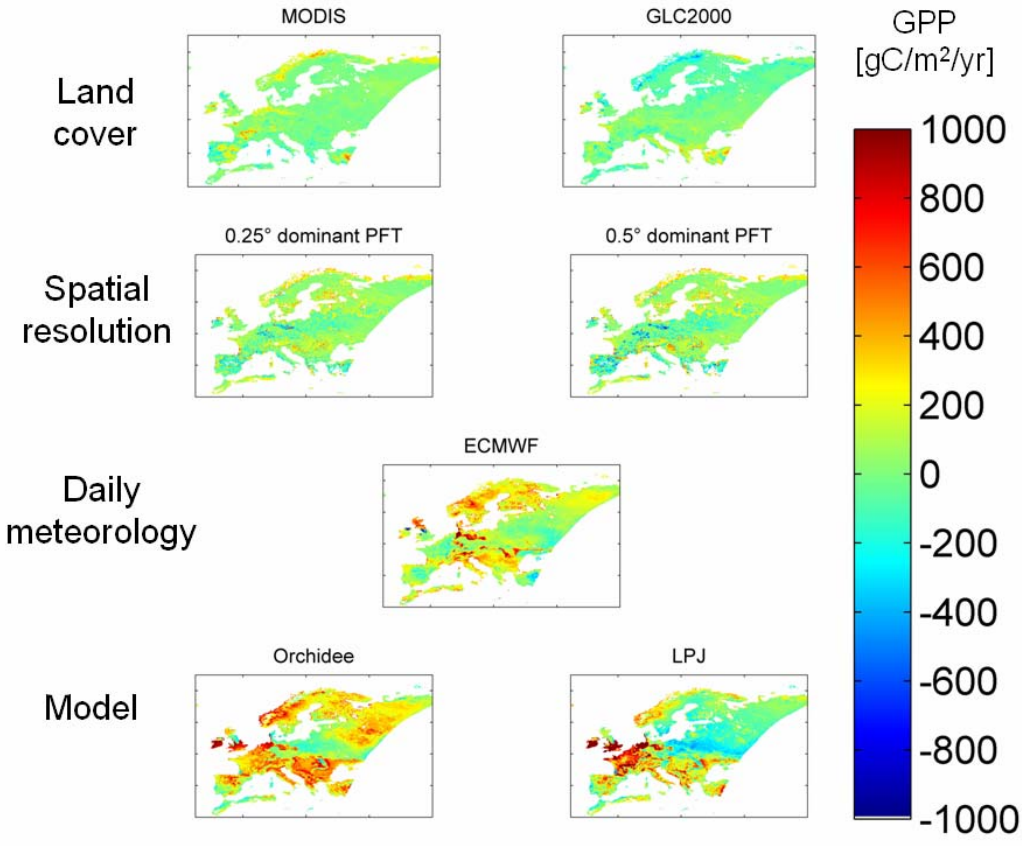


Figure 3-2: Difference maps of mean European GPP 1981-2000 for alternative realisations (AR-REF).

Regarding the correspondence of interannual variations of GPP between the reference and alternative realisations we find the same general pattern: poor agreement when changing the meteorological forcing or the model (Figure 3-3). The ECMWF scenario shows almost no correlation of interannual GPP variations with the reference in large parts of Eastern Europe and the Mediterranean. When using different models, there are only small areas in north and north Eastern Europe where there is moderate to high correlation with the reference. In general the spatial pattern of unexplained temporal variance is similar for the LPJ and Orchidee simulations. This might imply that the Biome-BGC interannual pattern differs substantially from LPJ and Orchidee while the latter two may be similar. When correlating the interannual variations of LPJ with Orchidee the large disagreement in temporal variation decreases from on average 60 - 63% (with LPJ and Orchidee respectively) to 43% unexplained variance (figure not shown). The correlations improve for the boreal region but remain weak over the mid-latitude cropland belt and southern Europe (see Figure 3-7 in auxiliary material).

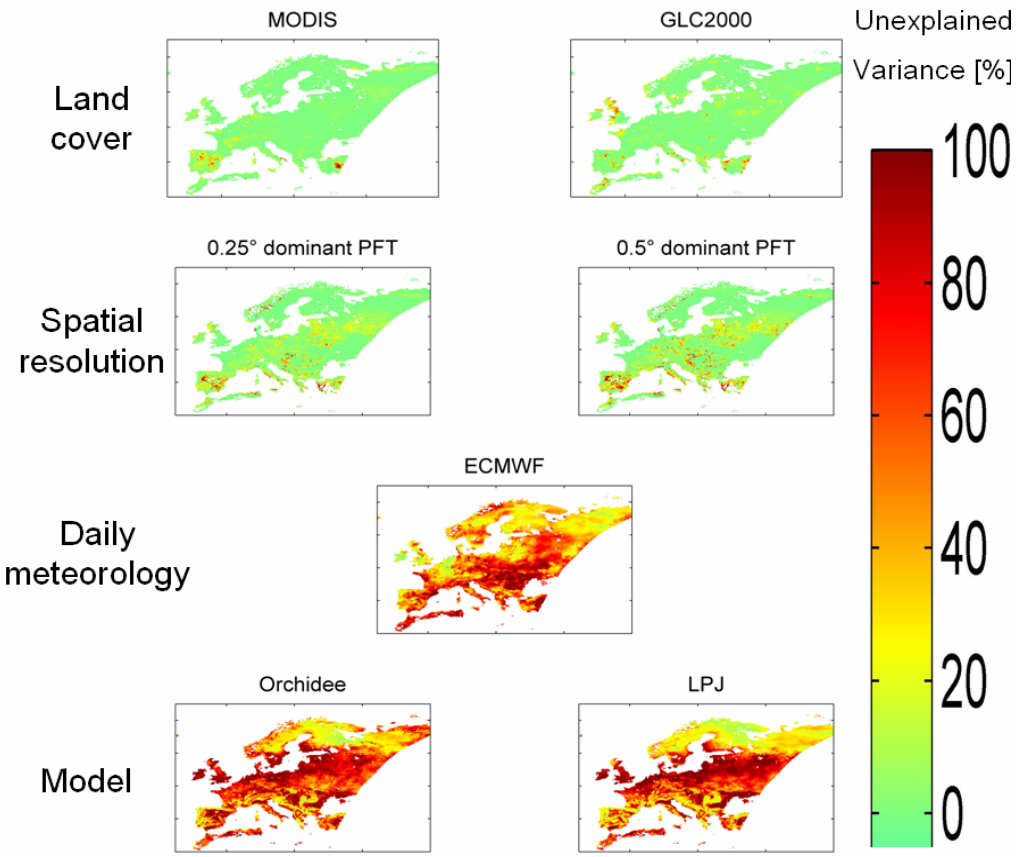


Figure 3-3: Effect of alternative realisations on the interannual variation of GPP. The fraction of variance that is not explained by the correlation R2 with the reference set up is shown for each pixel.

Figure 3-4 summarizes the effects of the different input data sets and models on GPP simulations. There is a clear hierarchy of uncertainties recognizable with a small effect of using different land cover maps, a somewhat higher but still relatively small effect of the spatial land cover resolution, a substantial effect due to changing the meteorological forcing, and the largest effect caused by using different models. The next sections discuss the individual factors in more detail.

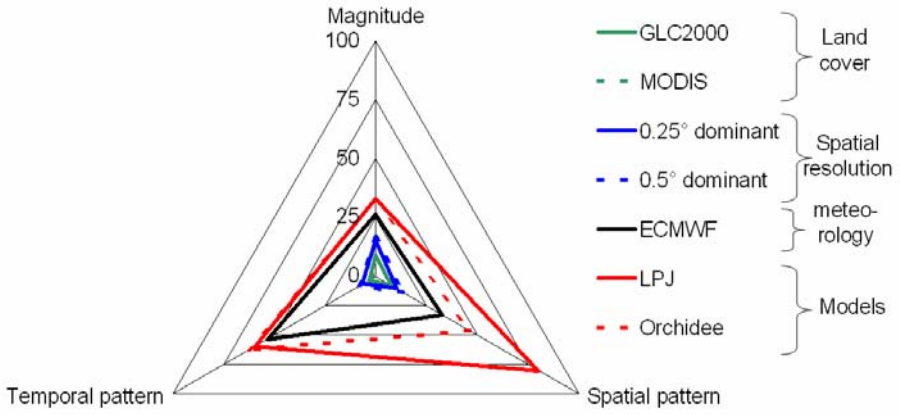


Figure 3-4: Effects of different model set-ups (alternative realisations) on the magnitude, spatial, and temporal pattern on GPP simulations over Europe. The measures are in % and based on the reference period 1981-2000 as explained in section 3.3.2. No difference to the reference set-up would be represented by the center where the axes intersect.

3.4.2 Land cover

We note that the land cover dataset effect is the smallest one for all investigated scenarios, not reaching 10% on neither magnitude, nor spatial or temporal pattern of modelled GPP. This coincides with findings of Beer, 2005 emphasizing the importance of land cover data to be included in carbon modelling but with small effects if different types of existing maps are used. Similar results are reported by Knorr and Heimann (2001) who found a rather small effect of changing the land cover data set on global NPP using the BETHY vegetation model.

Previous studies showed that various land cover classifications derived form remote sensing products have discrepancies among them, particularly in heterogeneous landscapes (Giri et al., 2005; Herold et al., 2006; Jung et al., 2006). Known global uncertainties for 1 km land cover datasets are in the order of 68 % area weighted overall accuracy considering all classes (Mayaux et al., 2006; Scepan, 1999). However, the map's uncertainty decreases if classes are aggregated

to PFTs and the larger grid sizes of the models (here 0.25° fractional). In addition, land cover types derived from satellite data represent direct and consistent spatial observations. The other investigated factors involve modelling and, thus, may contain larger error margins; at least from a theoretical point of view.

While simulations of GPP by TBM seem not to be very sensitive to the land cover map, we expect a much stronger effect on carbon stocks. Deviating cartographic standards and definitions lead to different forest extents and thus carbon stocks. Moreover, our conclusion of small effects of different land cover maps on simulated GPP is restricted to this class of models and data sets, which do not distinguish between crop functional types. To provide a substantial added value of future land cover products the remote sensing community needs to foster the separation of major crop types and management regimes (e.g. irrigated and non-irrigated). Implementing and improving the agricultural sector in biosphere models is currently a field of intensive research but partly hampered by the availability of adequate data sets.

3.4.3 Spatial resolution of the land cover map

We find the spatial resolution effect on the magnitude of GPP to be 15 % and 16 % for 0.25° and 0.5° dominant respectively. In terms of the spatial pattern, only 10 % and 14 % of the spatial variance remains unexplained. The temporal correlations are only minimal affected (max. 8% of unexplained variance). The fact that carbon flux calculations are to some extent sensitive to the pixel size have been shown previously and is consistent with this study (e.g. Kimball et al., 1999; Turner et al., 2000). Turner et al., 2000 used land cover maps of different spatial resolutions (from 25m to 1000m) to scale up field measurements from the north-western US and found that the difference between 25m resolution and 1000m resolution is ~12% for NPP. Kimball et al., 1999 run Biome-BGC with different spatial land cover resolutions over parts of the BOREAS region and found that NPP is affected by 2-14% by spatial aggregation effects.

This study indicates a more prominent effect of changing the spatial resolution compared to changing the land cover dataset. It is obvious that a 0.25-0.5 degree cell can only provide a rather coarse representation of the terrestrial vegetation heterogeneity if only the dominant type is mapped. Even the fractional PFT representation from a 1 km resolution land cover map may still

introduce representation bias in the carbon budget calculations. Representing vegetation at coarser pixel resolutions often leads to the suppression of certain types that can be important in terms of carbon cycling. In Europe, for example, this effect applies to the extensive agricultural areas and managed landscapes. Many trees and shrubs along field boundaries, roads, within cities as well as smaller patches of trees can be 'lost' in pixels that are mapped as e.g. crop because this dominates the 1 km mixed pixel. Such bias will soon be reduced by higher resolution global land cover data sets such as GLOBCOVER.

3.4.4 Daily meteorology

The model outputs are more affected by changing the meteorological drivers than for different land cover and spatial resolution options. Total GPP over the European domain of the ECMWF run is 20% higher than the simulations using REMO; the mean absolute difference over all grid cells being 26%. This order of magnitude is comparable to the study of Zhao et al., 2006 on the effect of different meteorological reanalysis (DAO, NCEP, ECMWF) on global GPP and NPP from the diagnostic MOD17 model. In their study, the largest differences occurred between NCEP and ECMWF with ~ 23 Gt/yr difference for GPP and even higher discrepancies for NPP (~27Gt/yr). Compared to model runs using meteorological observations, the relative error for GPP ranged from 16% (ECMWF) to 24% (NCEP); for NPP from 45% to 73%. Zhao et al., 2006 concluded that ECMWF appeared to perform best among the reanalysis data sets.

By investigating the differences of mean annual spatial fields of ECMWF and REMO (see Figure 3-8 in auxiliary material) we can explain the difference in the spatial patterns of GPP. Northern Europe is warmer and receives more radiation according to ECMWF which results in larger productivity, given that this area is expected to be primarily limited by radiation and temperature. The coinciding higher VPD seems not to counteract this effect suggesting little water limitation over the area in the model. Enhanced gross carbon uptake in southern Europe in the ECMWF runs is related to the higher rainfall in combination with lower VPD since water deficit controls photosynthesis to a large degree here.

We find the interannual variations of GPP due to the different meteorological driver data sets particularly striking. The temporal correlation between REMO and ECMWF radiation data is

very weak across almost entire Europe (see Figure 3-9 in auxiliary material) and it likely explains the differences in GPP interannual variability over Northern Europe where temperatures are highly correlated. Large discrepancies of interannual variations of radiation data sets have also been found by Hicke, 2005 who analysed the effect of using different radiation data sets (NCEP, GISS) on global NPP simulations from the CASA model. The author found only a small effect on total global NPP but large effects regarding the spatial pattern and especially interannual variations. For central and eastern Europe the large disagreement of GPP variations between REMO and ECMWF seems to originate from joint effects of differences in radiation, precipitation, and VPD, and likely nonlinear responses due interactions with nitrogen dynamics in the model (see section 3.4.5). The temporal correlations of the different data sets for all four meteorological variables are very low for southern Europe and all likely contribute to the deviations in simulated interannual GPP variations.

An in-depth analysis on the differences of the meteorological data sets and their origins would be insightful but is beyond the scope of this study. Cloud and aerosol physics that govern precipitation and radiation transfer is most likely the major factor that drives the differences among meteorological reanalysis. Orographic effects may further be important; certainly for mountainous regions which is visible in the difference of mean temperatures (see Figure 3-8 in auxiliary material) where REMO temperatures are substantially lower in the mountains due to its finer representation of topography. A detailed comparison and evaluation of REMO, ECMWF and also other possible meteorological model forcings (NCEP and CRU) is currently in progress (Chen et al, in preparation).

Important implications of our findings are that modelling studies focusing on interannual variations of carbon fluxes need to consider uncertainties in the meteorological forcing in their interpretations, especially exercises that aim to investigate effects of drought. In addition, it seems crucial to use the same meteorological drivers in model-intercomparison studies. Improved reanalysis would reduce uncertainties in the future if long term consistent time series are provided.

3.4.5 Biosphere models

Several model-intercomparison studies have shown substantial differences among models (e.g. Cramer et al., 1999; Roxburgh et al., 2004) while mechanistic explanations for the differences have been rarely presented. Such task is difficult given that models differ in many respects and isolating the effect of certain alternative parameterisations is hardly possible given the interactions within the model. We aim to infer the likely most important causes of model differences here to guide future modelling studies, which will allow a more objective judgement on the degree of realism and robustness.

Spatial patterns

Key factors that likely cause the major differences are related to the model representation of the agricultural sector, nitrogen dynamics, soil hydrology, parameter values, and sensitivity to meteorological conditions, the latter being partly linked to the former factors. LPJ is the only model in this study that has a realistic representation of the agricultural sector. Biome-BGC and Orchidee represent crops as productive natural grassland assuming fertilization (Biome-BGC) or enhanced photosynthetic capacity (Orchidee). The large disagreement among the models in terms of mean annual GPP patterns in the cropland regions is certainly related to this issue (see Figure 3-10 in auxiliary material).

Among the three models, nitrogen limitation is only accounted for explicitly in Biome-BGC. This is expected to result in differences among the models along gradients of nitrogen availability such as the transition from boreal to temperate ecosystems. In a recent study we investigated how well the three same models reproduce the spatial gradient of GPP of forest ecosystems across Europe (Jung et al., 2007a). The models appeared to produce a too weak gradient from boreal to temperate forests. We inferred that this resulted primarily from simulating almost no change of LAI, and thus light absorption in the case of LPJ and Orchidee. Biome-BGC performed somewhat better here indicating the effect of increasing nitrogen availability on LAI and light harvesting. GPP is particularly sensitive to the simulated LAI in the range 0 to 3. GPP becomes insensitive to LAI variations when LAI exceeds a value of 4 because changes in light interception become marginal. The significance of the role of nitrogen has also been recently emphasized by Magnani et al., 2007 who suggested that observed relationships between forest GPP and mean

annual temperature (e.g. Reichstein et al., 2007b) are strongly related to a corresponding gradient of nitrogen availability.

Parameter sensitivity studies (White et al., 2000; Zaehle et al., 2005) have also pointed to the significance of those related to LAI and light absorption such as light extinction coefficient and specific leaf area. Parameters related to maximum photosynthetic capacity and stomata conductance appeared to be at least equally sensitive. A whole series of parameters is associated with PFTs leading to an imprint in the spatial pattern of the simulations according to the PFT distribution while spatial variations within PFTs may be underestimated. Such spatial imprint is visible when one compares the distribution of prescribed vegetation types and simulations of GPP and maximum LAI (see Figure 3-10 in auxiliary material). A better understanding of variations and covariations of sensitive parameters in the future may allow removing some of the constrains by fixed parameters and more confidence in predictions. Recent studies link the coordination of plant traits (e.g. Wright et al., 2004a) to optimisation principles in ecosystems and this approach represents possibly an avenue to overcome some of the limitations (Anten, 2002; Anten, 2005; Hikosaka, 2005; Shipley et al., 2006).

Interannual variability

The low correspondence of simulated interannual variations of Biome-BGC with LPJ and Orchidee is striking. We can gain a first insight into the principal mechanism of GPP variability within the models by decomposing GPP into its 'biophysical' and 'ecophysiological' component, absorbed photosynthetic active radiation (APAR) and radiation use efficiency (RUE) respectively (Figure 3-5). The spatial pattern of the strength of interannual GPP variation partly differs among models. Biome-BGC and Orchidee show larger variability than LPJ in southern England, the North Sea cost and parts of France while LPJ generates larger variability on the Iberian peninsular and east of the Adriatic Sea than the other two models. Biome-BGC predicts lower variability north of the Black Sea than LPJ and Orchidee. In general the variation of RUE is stronger then the variation of APAR although differences among models are apparent too. LPJ shows smallest, Biome-BGC intermediate, and Orchidee largest variation of APAR. The relatively higher APAR variability of Biome-BGC and Orchidee result partly from the lower mean maximum LAI (see Figure 3-10 in auxiliary material) in the range where fAPAR is sensitive to variations of LAI. In addition carbon allocation operates on a daily time step in

Biome-BGC and Orchidee and therefore allows for greater variability of the leaf carbon pool. LPJ in contrast, has annual allocation and leafs are shed only at the end of a season for deciduous vegetation. Variations of fAPAR in the models are somewhat both, cause and consequence of GPP since LAI depends on NPP. Corroboration against APAR data calculated from remotely sensed fAPAR (Gobron et al., 2006) suggests that the variation of APAR may be overestimated by Orchidee in the case of crops and broadleaf trees, by Biome-BGC in the case of broadleaf trees while LPJ may produce too little APAR variability in general (see Figure 3-11 in auxiliary material). However, given that RUE varies more and its variations are more strongly correlated with the variations of GPP (see Figure 3-12 in auxiliary material) reveals a dominant ecophysiological control of GPP interannual variability in the models. This is consistent with ongoing studies from Reichstein et al (in prep) for forest ecosystems in Europe. GPP and RUE variations as well as their differences among models are predominant in the mid and low latitudes of Europe suggesting that model differences may result primarily from water stress effects. Since RUE lumps a number of different model components as well as their interactions into a single number we further investigate the relationship and sensitivity of modelled GPP to meteorological conditions.

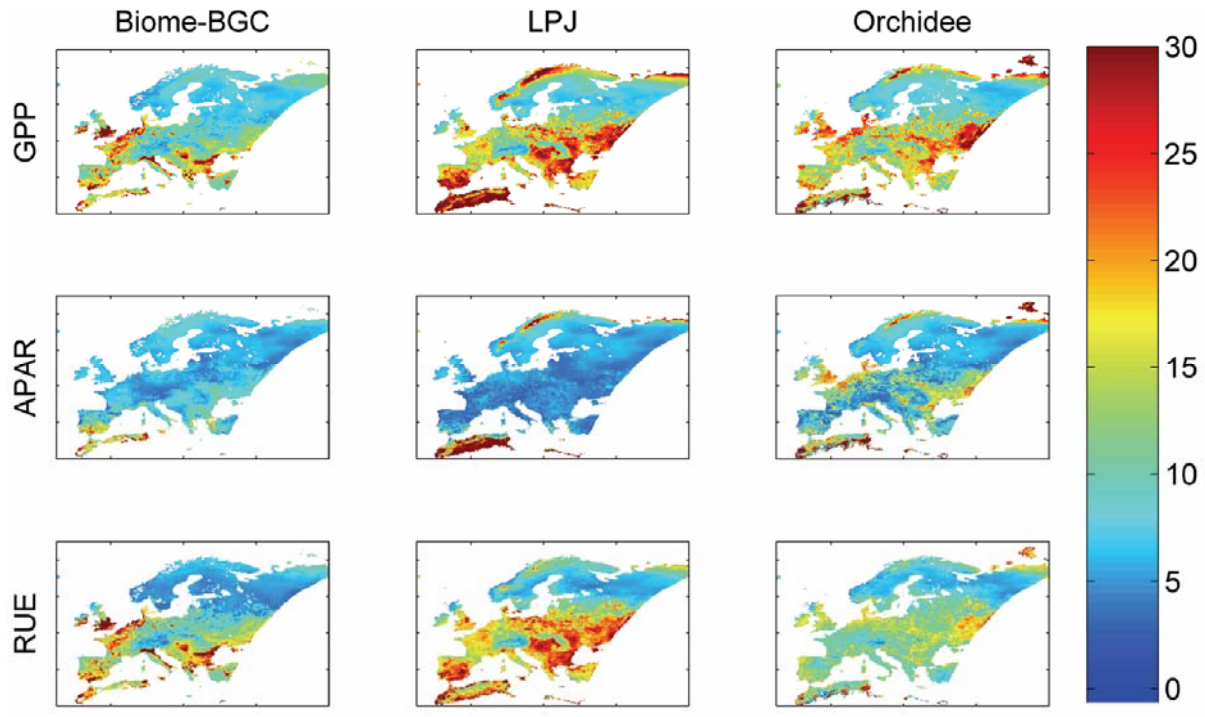


Figure 3-5: Coefficient of variation (standard deviation divided by mean, in %) of GPP, APAR, and RUE for Biome-BGC, LPJ, and Orchidee (1981-2000). The variation of a product (GPP) is predominantly controlled by the factor (APAR or RUE) that shows larger variability. The figure reveals predominant ecophysiological (RUE) control of interannual variability of GPP in the models.

The first principal component (PCA1) explains 84 % of the variation of the meteorological data set (Table 3-2). The different meteorological variables contribute to roughly the same amount to this axis as can be seen from the eigenvectors; negative values are associated with high radiation, temperature and VPD but low rainfall, positive values the opposite. PCA1 represents a typical weather gradient from 'warm, sunny, and dry' to 'cool, cloudy, and moist'. The three models show strong negative correlations with PCA1 over northern Europe, i.e. summer GPP increases correlate with temperature and radiation increases (Figure 3-6). The sensitivity of the models, expressed as the slope of the regression line, is similar and relatively small as is the GPP variability over this area from the models (see Figure 3-5).

Principal component axis	Variance explained [%]	Eigenvectors			
		Radiation	Temperature	VPD	Precipitation
PCA1	84	-0.283	-0.280	-0.283	0.241
PCA2	11	-0.239	-0.597	-0.311	-1.340
PCA3	3	2.151	-0.234	-1.798	0.137
PCA4	2	1.337	-2.889	2.020	0.581

Table 3-2: Result of the principal component analysis (PCA) of the meteorological input data. The PCA was performed on z-score standardized mean data from June to August for each year (mean removed). The eigenvectors give the contribution of the meteorological variables to the different principal component axis.

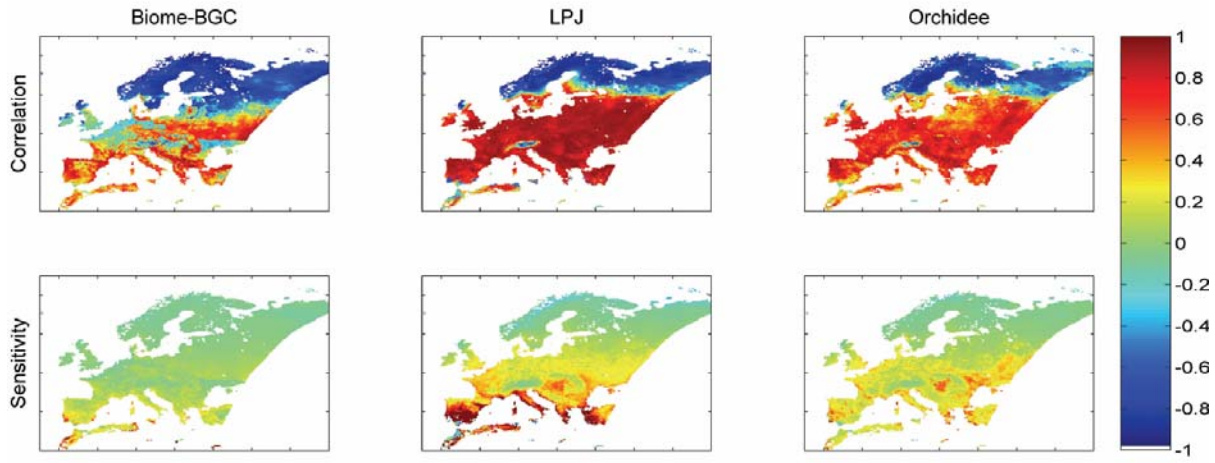


Figure 3-6: Correlation and sensitivity (slope of regression line) of relative GPP variations to the first principal component of mean JJA meteorology. Positive correlations mean that GPP increases with temperature, radiation, and VPD and decreases with rainfall (Northern Europe); negative correlations the opposite (Central and southern Europe). It shows that the relationship between summer meteorology and simulated GPP is partly different for Biome-BGC and that the three models differ in their sensitivity to meteorological conditions.

For the mid and low latitudes of Europe, the relationship reverses, i.e. simulated GPP correlates positively with rainfall and negatively with radiation, temperature and VPD. Variations of moisture appear to drive variations of GPP here. The transition from temperature and radiation control to moisture control of GPP is slightly further south in Biome-BGC. LPJ and Orchidee have similar spatial correlation patterns with PCA1, showing a ubiquitous relationship with moisture while the relationship is stronger for LPJ. Interestingly, Biome-BGC shows no relationship with PCA1 in large parts of the European mid-latitudes, particularly in the maritime parts of Western Europe. Photosynthesis in Biome-BGC does apparently not always respond to moisture variations in summer. This effect originates most likely from interactions with the nitrogen cycle in Biome-BGC. In years when meteorological conditions would allow high levels of productivity this level cannot be reached because the nitrogen demand exceeds the supply. Biome-BGC calculates the nitrogen demand, based on predefined C:N ratios of different structural compartments of the vegetation, and if the supply is insufficient, the amount of carbon assimilated is corrected down to the level where it matches the nitrogen supply. Productivity, leaf turnover and decomposition, being itself controlled by temperature, soil moisture and nitrogen, determine nitrogen supply. In accordance to our findings, Kirschbaum et al., 2003 have shown that the feedbacks between the carbon and nitrogen cycle in the CenW model have substantial impact on interannual variations of NPP and NEP in Australia. The interactions of carbon and nitrogen dynamics can lead to complex patterns that are often not simply related to meteorological conditions of a growing season. We can partly attribute the substantial disagreement of the interannual variations of GPP between Biome-BGC with LPJ and Orchidee to interactions with nitrogen in Biome-BGC. This feedback between above ground productivity and decomposition in the soil deserves further attention in the future since it has a large effect in the model that includes a nitrogen cycle. For natural ecosystems, Anten, 2005; Hikosaka, 2005 have shown that interactions with the nitrogen cycle shape ecosystem traits that control photosynthesis assuming optimisation principals in ecosystems. Such approach may further be considered in the context of global modelling aiming to predict, rather than prescribe, sensitive ecosystem properties.

The sensitivity of the different models to moisture variations is substantially different. Biome-BGC shows least sensitivity and LPJ greatest sensitivity (Figure 3-6). Orchidee displays only slightly larger sensitivity than LPJ in parts of Eastern Europe. Several structural model

components play particularly important roles in determining the response to variations of moisture: (1) interactions with the nitrogen cycle in the case of Biome-BGC as discussed above, (2) the representation of the soil environment, (3) canopy conductance and its feedback to photosynthesis and soil moisture, and (4) direct water stress effects on photosynthetic capacity.

A smaller sensitivity of Biome-BGC to water stress can be expected given that it represents the soil as a simple one layer bucket without accounting for a differentiated root profile on plant available water. LPJ and Orchidee use two layer models with particular root profiles and depths, depending on the vegetation type. LPJ has a fixed depth of the upper layer of 50 cm while Orchidee's upper layer has dynamic depth, which represents the zone below field capacity. Drying of the upper layer with a higher concentration of roots there makes the two models more sensitive to water stress than Biome-BGC, particularly for herbaceous vegetation with short rooting depths. Orchidee is the only model among the three that uses a parameterisation of soil water stress on photosynthetic capacity (Vcmax).

The central linkage between the water and carbon cycle is canopy conductance, which determines intercellular CO_2 concentrations available for photosynthesis and water loss through transpiration, and differences among models in this respect are likely critical. Biome-BGC uses are Jarvis type of approach where a predefined maximum canopy conductance is reduced in a multiplicative scheme of scalars according to environmental conditions (VPD, soil moisture, temperature, radiation, and nitrogen availability). Canopy conductance affects photosynthesis but not the other way round and the feedback comes from the depletion of soil water. In LPJ, canopy conductance, photosynthesis and transpiration are intimately linked. The equations are solved iteratively to yield consistent results according to water demand, and supply from the soil. The strong connection to the soil water status causes downregulation of canopy conductance and photosynthesis as to not fully deplete soil water storage. This mechanism is likely responsible for the strong sensitivity of LPJ to water availability. Orchidee uses the Ball-Berry formulation that relates canopy conductance to assimilation and air humidity and the respective equations are solved iteratively, thus representing a two way interaction between canopy conductance and assimilation as in LPJ. In contrast to LPJ, canopy conductance in Orchidee is sensitive to air humidity rather than to soil moisture. Differences of sensitivity between LPJ and Orchidee as depicted in Figure 3-6 may well be related to this factor.

3.5 Conclusions

We have presented a systematic study on how alternatives of the model set-up affect magnitude, spatial, and temporal patterns of GPP simulations over Europe, using different land cover maps, spatial land cover resolutions, meteorological data sets, and process oriented TBMs. We found a clear hierarchy of effects: a small effect of using different land cover maps, a somewhat higher but still relatively small effect of the spatial land cover resolution, a substantial effect due to changing the meteorological forcing, and the largest effect caused by using different models.

Differences in the meteorological model forcings affect particularly interannual variations of modelled GPP. Carbon cycle modelling studies that focus on interannual variations need to consider these uncertainties. Furthermore, we strongly recommend using the same meteorological driver data set for each model in intercomparison studies, since otherwise it is not possible to differentiate between model and driver effect when comparing the simulations.

From a model structure point of view, differences between the models in terms of simulating interannual variations of gross carbon uptake are strongly linked to the way of how and if biogeochemical cycles (carbon, water, and nitrogen) interact within the models which controls their sensitivity to meteorological conditions. The related mechanisms used in the models should be clarified and verified since these may shape the carbon cycle climate feedback in Earth system models. We highlight the effect of carbon-nitrogen interactions in altering the effect of interannual climate variability on carbon flux variations, here GPP. Water stress effects impact on photosynthesis differently in the models. We suggest revisiting formulations of canopy conductance which represents the central linkages of the carbon and water cycle in the models. In general the representation of soil environment in the models deserves particular attention since processes controlling water and nutrient availability operate here. A sound representation of ecosystem functioning is necessary to capitalize on recent concepts of ecosystem dynamics to changing environmental conditions such as reorganisations of traits to maximize resource use efficiency. Such approaches may lead to more confidence in large scale modelling, both spatially and temporally, while substantial research still needs to be done in this respect.

3.6 Auxiliary material

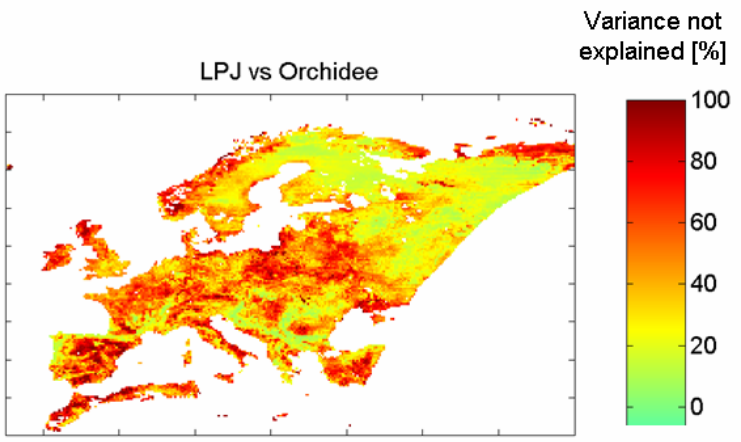


Figure 3-7: The fraction of interannual variance that is not explained by the correlation R^2 between LPJ and Orchidee for each pixel.

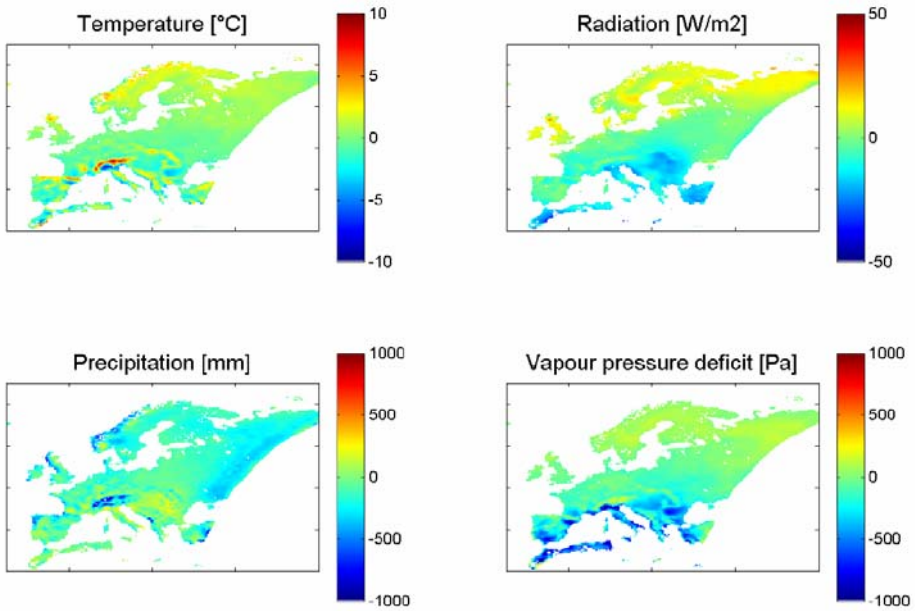


Figure 3-8: Difference of mean annual (1981-2000) meteorological variables of ECMWF and REMO (ECMWF-REMO).

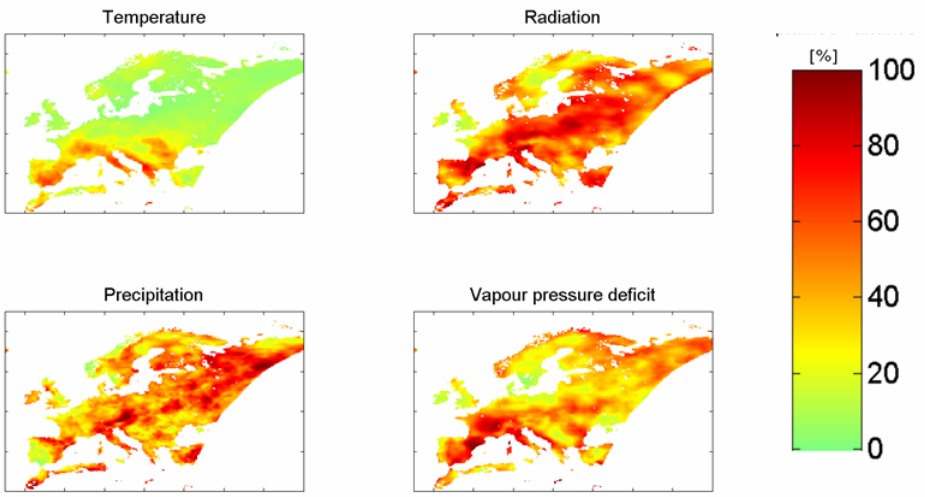


Figure 3-9: The fraction of variance that is not explained by the correlation R^2 ($1-R^2$) between meteorological forcing fields between ECMWF and REMO for each pixel.

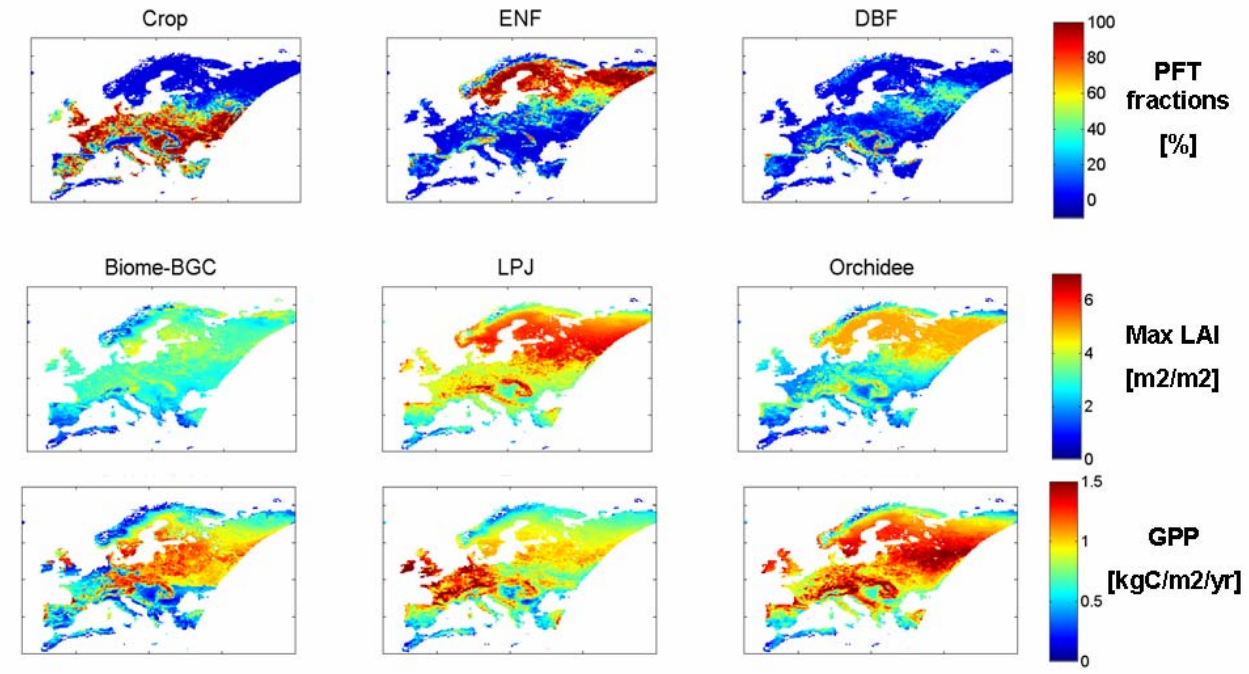


Figure 3-10: Fractions of the most important vegetation types used as input and mean maximum LAI and annual GPP (1981-2000) as simulated by Biome-BGC, LPJ, and Orchidee. The comparison between the distribution of PFTs and the simulations reveals partly spatial imprints as a consequence of PFT specific parameters.

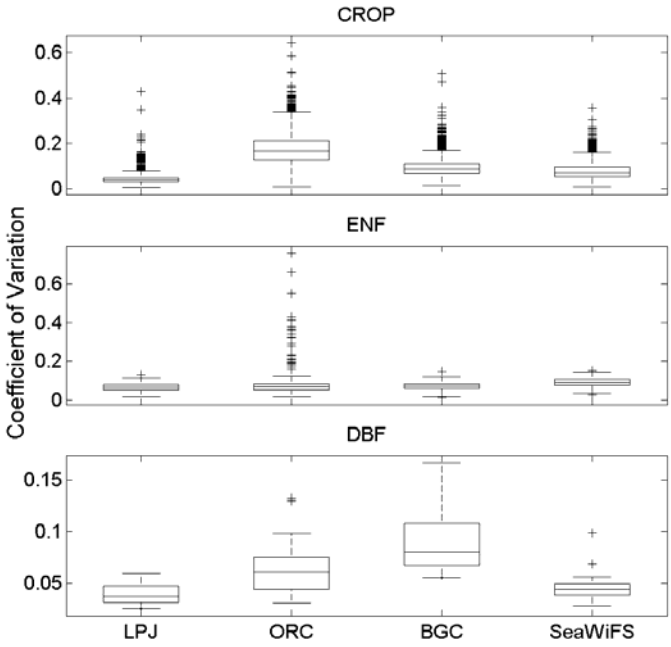


Figure 3-11: Comparison of the variation of APAR for different models and an independent estimate calculated from the SeaWiFS FAPAR product. The box marks the interquartile range with the median within it; crosses mark outliers defined as data points that are beyond median +/- 1.5 * interquratile range. The 'whiskers' give the range of the data but maximal data upto median +/- 1.5 * interquartile range.

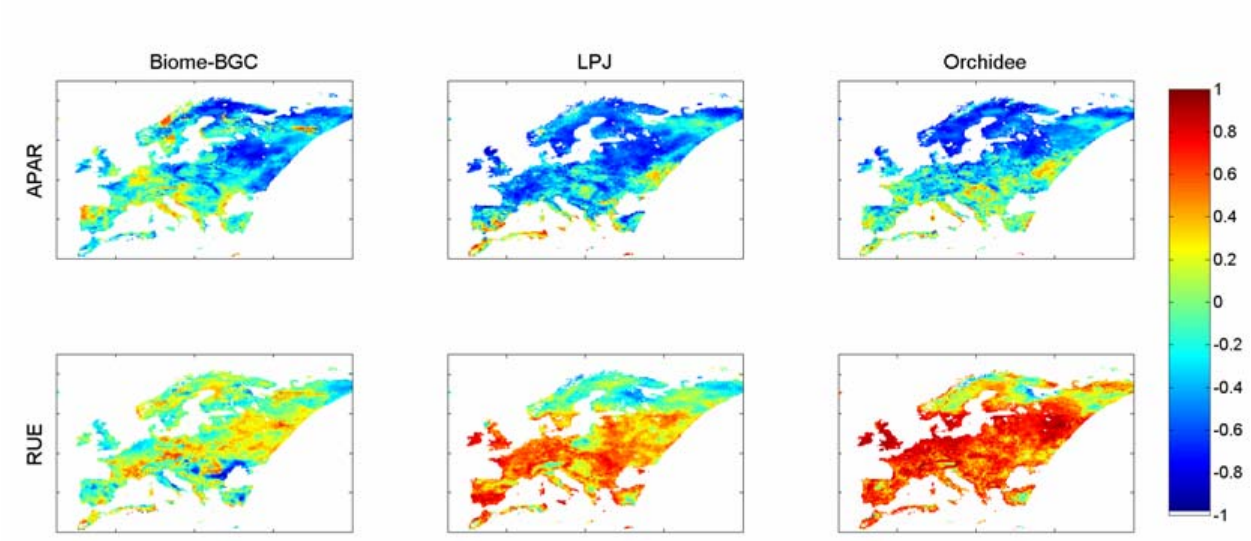


Figure 3-12: Maps of Pearson's correlation coefficient between GPP and APAR and RUE (1981-2000). The correlations are not very meaningful in regions of little variability (see Figure 3-5 in the paper). Negative correlations between APAR and GPP result from water stress which often covaries with radiation.

4 Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe[3]

Abstract

Three terrestrial biosphere models (LPJ, Orchidee, Biome-BGC) were evaluated with respect to their ability to simulate large-scale climate related trends in gross primary production (GPP) across European forests. Simulated GPP and leaf area index (LAI) were compared with GPP estimates based on flux separated eddy covariance measurements of net ecosystem exchange and LAI measurements along a temperature gradient ranging from the boreal to the Mediterranean region. The three models capture qualitatively the pattern suggested by the site data: an increase in GPP from boreal to temperate and a subsequent decline from temperate to Mediterranean climates. The models consistently predict higher GPP for boreal and lower GPP for Mediterranean forests. Based on a decomposition of GPP into absorbed photosynthetic active radiation (APAR) and radiation use efficiency (RUE), the overestimation of GPP for the boreal coniferous forests appears to be primarily related to too high simulated LAI - and thus light absorption (APAR) – rather than too high radiation use efficiency. We cannot attribute the tendency of the models to underestimate GPP in the water limited region to model structural deficiencies with confidence. A likely dry bias of the input meteorological data in southern Europe may create this pattern.

On average, the models compare similarly well to the site GPP data (RMSE of ~30% or 420 gC/m2/yr) but differences are apparent for different ecosystem types. In terms of absolute values, we find the agreement between site based GPP estimates and simulations acceptable when we consider uncertainties about the accuracy in model drivers, a potential representation bias of the eddy covariance sites, and uncertainties related to the method of deriving GPP from eddy covariance measurements data. Continental to global data-model comparison studies should be fostered in the future since they are necessary to identify consistent model bias along environmental gradients.

[3] Published as: Jung, M., Le Maire, G., Zaehle, S., Luyssaert, S., Vetter, M., Churkina, G., Ciais, P., Viovy, N., Reichstein, M. (2007): Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe. *Biogeosciences*, 4, 647-656

4.1 Introduction

Continental to global scale simulations of the land carbon cycle are subject to uncertainties related to model structure, parameters, and input driver data (McGuire et al., 2001; Moorcroft, 2006; Morales et al., 2005; Zaehle et al., 2005). Confronting simulations with measurements allows assessing the model's performance, gaining confidence in the model predictions and/or identify major issues with the model structure. Such comparisons have been repeatedly made for single or few intensively investigated eddy covariance flux measurement sites when it was possible to parameterise and drive the models with in-situ data (e.g. Churkina et al., 2003; Kucharik et al., 2006; Morales et al., 2005). These analyses revealed important insights regarding the credibility of the model's dynamics and simulated temporal variations. However, models designed for the continental to global scale should also be evaluated on that scale, i.e. investigating how well the broad patterns along large environmental gradients are reproduced. Such studies have rarely been presented, primarily due to a lack of consistent synthesis work of carbon flux measurements. Global data for net primary productivity (NPP) are available (Scurlock et al., 1999, http://www-eosdis.ornl.gov/NPP/npp_home.html) but prove to be difficult to use as benchmarks (e.g. Cramer et al., 1999; Zaehle et al., 2005). Because compilations of NPP measurements suffer from inconsistent methodologies, individual values from different sites and investigators are often not compatible (but see Luyssaert et al., accepted). In addition, NPP data are known to be biased low to an unknown extent and there is strong indication that this bias can change substantially for different climate regions (Luyssaert et al., accepted).

Consistent estimates of gross primary production (GPP) are now becoming available from the eddy covariance measurement community based on methods that separate measured net ecosystem exchange (NEE) into GPP and ecosystem respiration (Reichstein et al., 2005). In this study we evaluate simulated GPP from three global biogeochemical models (LPJ, Orchidee, Biome-BGC) for forest ecosystems in Europe. Our study is consistent with, and complements a recent model intercomparison project within the Carboeurope-IP project that aims to understand, quantify, and reduce uncertainties of the European carbon budget (http://www.carboeurope.org/). We investigate the performance of the models to reproduce the broad pattern suggested by eddy covariance based GPP along a mean annual temperature gradient running from the boreal to the Mediterranean. We evaluate to what extent we can be confident with European scale simulations

of forest GPP, and aim to identify consistent patterns of correspondence and mismatch with the data. We further propose a simple method of decomposing GPP into APAR and RUE that aids in the diagnoses of model performance using ancillary leaf area index (LAI) measurements.

4.2. Materials and Methods

2.1 Site data

The observational site data we use originate from the recent data compilation of Luyssaert et al., accepted. We extracted all available data from sites with GPP (annual sums) or LAI measurements (annual maximum) for Europe. We excluded sites from mixed forests (mixed plant functional types or PFTs), manipulative experiments where the forest was fertilized or irrigated, as well as recently disturbed plots and clear cuts. Finally, 37 and 47 sites for GPP and LAI respectively are available of which 22 have both GPP and LAI estimates (Figure 4-1).

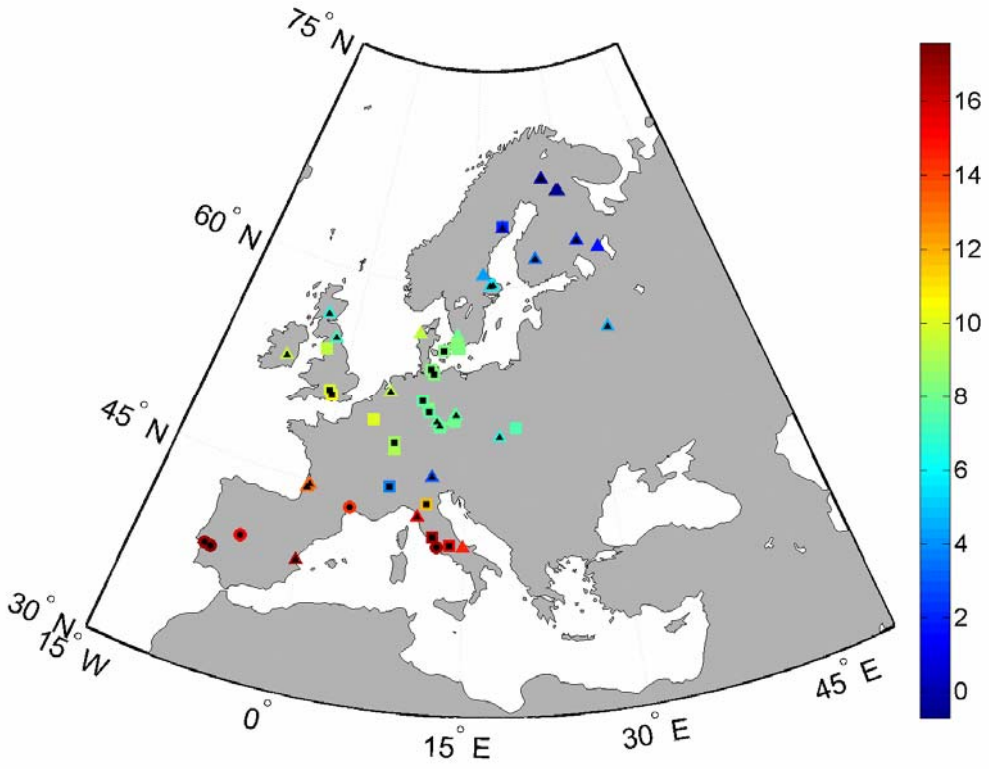


Figure 4-1: Spatial distribution of GPP and LAI measurements. Sites with GPP measurements have a black filling. Triangles: evergreen needleleaf forests, squares: deciduous broadleaf forests, circles: evergreen broadleaf forests. Colour represents mean annual temperature in °C (1981-2000 mean from REMO).

The GPP data originate from Carboeurope eddy covariance tower sites that measure the net ecosystem CO_2 exchange (NEE). The data represent the time period from approximately 1996 to 2005 with a bias towards recent times. The NEE fluxes had been separated into GPP and ecosystem respiration (R_{eco}) by subtracting R_{eco}. R_{eco} had been calculated based on its night time temperature sensitivities, the vast majority according to Reichstein et al., 2005.

LAI measurements are partly based on different methods; indirect optical methods have been used primarily. By means of the Lambert-Beer's law, we converted LAI to the fraction of absorbed photosynthetic active radiation (fAPAR) which is the key variable for light absorption and thus GPP (Equation 4-1). The transformation of LAI to fAPAR allows a better interpretation to what extent a simulated mismatch in light harvesting might be responsible for the mismatch of simulated and observed GPP since light transmission is a negative exponential function of LAI. The Lambert-Beer's law as 1-D representation of canopy radiation transfer is also used in the three models to estimate light extinction.

$$fAPAR = 1 - e^{-k \times LAI} \qquad \text{(Equation 4-1)}$$

where k denotes the light extinction coefficient, assuming k = 0.5 for conifers and k = 0.58 for broadleaf trees. The conversion of LAI to fAPAR implies larger discrepancy of light harvesting at low LAI values and smaller discrepancy at high LAI values. For example, the fAPAR difference between LAIs of 2 (fAPAR ~ 0.63) and 4 (fAPAR ~ 0.86) is much larger than between LAIs of 6 (fAPAR ~ 0.95) and 8 (fAPAR ~ 0.98). While LAI can be considered to be a determinant of GPP in the range of 0 to 4 it becomes more a consequence of GPP beyond an LAI of 4 when changes of LAI have only minor effects on light absorption.

4.2.2 Model simulations

We performed simulations at the locations of the measurement sites using three state of the art global biogeochemical models: LPJ, Orchidee, and Biome-BGC. The models are described in detail in Sitch et al., 2003, Krinner et al., 2005, and Thornton, 1998; Thornton, 2002 respectively. We used the same input data for each model, according to a modelling protocol that is consistent with model intercomparison studies by Trusilova et al., in review and Jung et al., in review to

ensure comparability. We prescribed the PFT according to the prevailing vegetation type given in the database by Luyssaert et al., accepted. No site history was prescribed that accounts for age and management related effects; the models simulate mature forest stands. Soil water holding capacity and meteorological model drivers originate from gridded data sets with a spatial resolution of 0.25°. Water holding capacity data are based on IGBP-DIS, 2000 soil texture data. Meteorological model input from 1958-2005 is from a regional climate model (REMO, Jacob and Podzun, 1997) that was driven with NCEP reanalysis (Kalnay et al., 1996) at the boundaries of the European model domain (Feser et al., 2001). Details about model drivers and the modelling protocol are available in Trusilova et al., in review and the Carboeurope-IP model intercomparison homepage (http://www.bgc-jena.mpg.de/bgc-systems/projects/ce_i/index.shtml).

For consistency, we matched simulated GPP and LAI with the site data on a site by site and year by year basis. Subsequently, the yearly data were aggregated (averaged) to the site level. In cases two or more measurement sites with the same PFT fell within the same 0.25° gridcell (i.e. identical model output), data on site level were further averaged to gain more representative values on the 0.25° gridcell level.

4.2.3 Decomposing GPP into APAR and RUE

We decomposed GPP [gC m^{-2} yr^{-1}] into absorbed photosynthetic active radiation (APAR [MJ m^{-2} yr^{-1}) and radiation use efficiency (RUE [gC MJ^{-1}]). This procedure provides further information about possible causes of mismatch between simulated and site eddy covariance based GPP.

$$GPP = APAR \times RUE \qquad \text{(Equation 4-2)}$$

We calculate APAR for the models according to a standard method used in model intercomparisons from monthly mean leaf area index and radiation (e.g. Bondeau et al., 1999; Ruimy et al., 1999) (Equation 4-3). For the actual forest, there is commonly only one annual LAI measurements that represents approximately the annual maximum. In order estimate APAR for the forest sites we use the simulated seasonal pattern of fAPAR from the models but scale the simulated maximum fAPAR to the measured fAPAR (both calculated from LAI). In this way we calculate the APAR of the forest sites by using the modelled leaf phenology but correct for the

wrong magnitude of modelled fAPAR. Our approach yields consistent estimates of APAR for the simulated and actual forest that allows comparison among them.

$$APAR = \sum_{m=1}^{12} fAPAR_{sim\ m} \times PAR_m \left[\times CF\right] \qquad \text{(Equation 4-3)}$$

$$\text{with } CF = \frac{fAPAR_{obs}}{fAPAR_{\max\ sim}} \qquad \text{(Equation 4-4)}$$

Where, APAR denotes the absorbed photosynthetic active radiation [MJ m^{-2} yr^{-1}], m is an index for the month, fAPAR is the fraction of absorbed photosynthetic active radiation, calculated according to Equation 4-1, the subscript sim denotes the simulation, PAR is photosynthetic active radiation [MJ m^{-2} $month^{-1}$] from REMO, assuming PAR = 0.48 x global (short wave) radiation. CF is a correction factor that was only used for the estimation of APAR at the actual forest sites based on one LAI measurement.

The calculation is performed for all years with GPP measurements with subsequent averaging over the years. Since the seasonal pattern of simulated fAPAR (leaf phenology) may differ among models we calculate an actual site APAR for each model. The differences between the site APARs for different models are then entirely related to differently simulated phenology not due to the maximum reached LAI. Site and modelled RUE can now be calculated based on Equation 4-2, i.e. using eddy covariance flux separated GPP and site APAR, and simulated GPP and simulated APAR respectively.

Our method to decompose GPP into APAR and RUE for both, simulated and actual forest ecosystems uses several necessary simplifications and is only a first order approximation. We do not account for factors like albedo, diffuse radiation, and complex canopy structure that are relevant to the realistic estimations of fAPAR from LAI. Moreover, the models use internally partly different representations of the energy budget (e.g. albedo), differ slightly in the PFT specific light extinction coefficients and assumptions about upscaling of light absorption from tree to grid cell level. The derived absolute values of APAR and RUE are neither comparable among models nor to field measurements. However, our approach yields consistent results for

APAR and RUE between simulated and actual forest ecosystem, since we apply the same method. It is an efficient way of assessing whether systematic differences of light harvesting can explain the mismatch between observed and modelled GPP.

A drawback of the method is that it does not account for the observed seasonal pattern of light absorption due to a lack of measurement data with high temporal resolution. Consequently, we rely on the modelled seasonal pattern of LAI. Using the simulated seasonal pattern of LAI is only a minor issue for evergreen coniferous forests and we therefore restrict the application of the decomposition method to coniferous forests. Using the method for deciduous vegetation would require a priori confidence in the simulated timing of leaf onset, maximum LAI and leaf senescence for all three models. Alternatively, the availability of seasonally resolved measurements of LAI and/or of light interception for many sites would make it possible to use the actual observed seasonal cycle of leaf phenology.

4.3 Results and Discussion

4.3.1 Gross Primary Productivity

LPJ, Orchidee, and Biome-BGC reproduce the general pattern of GPP changes along the temperature (MAT) gradient across Europe. Across the continent GPP increase from boreal to temperate and subsequently decreases from temperate to Mediterranean regions (Figure 4-2). However, the models consistently predict higher GPP for the boreal and lower GPP for the Mediterranean zone than suggested by eddy covariance based GPP. Variations of GPP by the LPJ model are smaller than indicated by eddy covariance based GPP and the other two models Orchidee, and Biome-BGC. By comparing the means of observed and modelled GPP over all sites, we find that all the three models predict on average lower GPP than the eddy covariance based (Table 4-1), while the difference between simulated and observed means is not significant for Orchidee and Biome-BGC (according to a one-way analysis of variance (ANOVA)).

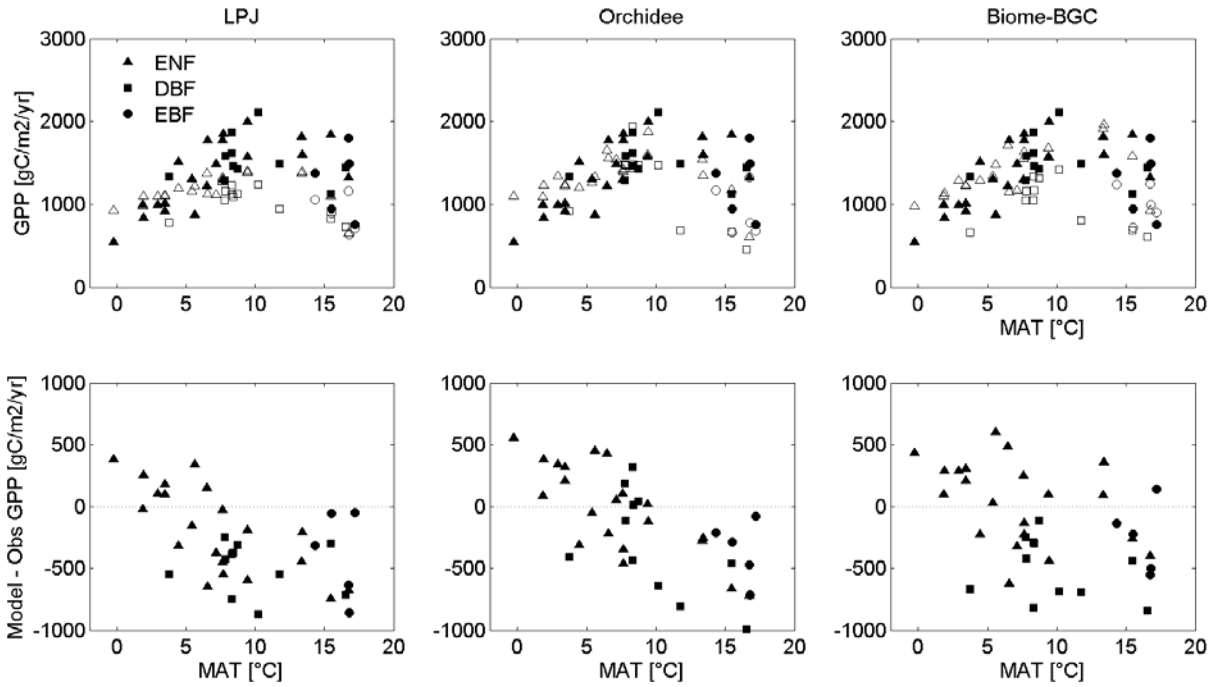


Figure 4-2: Top panel: eddy covariance flux separated (filled markers) and modelled (open markers) GPP along the mean annual temperature gradient across Europe. Bottom panel: difference between modelled and eddy covariance flux separated GPP along mean annual temperature (MAT, 1981-2000 mean based on the REMO data set). ENF: evergreen needleleaf forests, DBF: deciduous broadleaf forests, EBF: evergreen broadleaf forests.

The root mean square error of prediction (RMSE) over all sites is in the order of 420 gC/m2/yr (~ 30%) for the three models (Table 4-1). The stratification by ecosystem types reveals differences among models as well as among forest types and reveals individual contributions to the overall RMSE. On average, the RMSE is smallest for temperate coniferous sites (16-25 %) and largest for Mediterranean forest ecosystems (21-61%), which has also been observed by Morales et al., 2005 with respect to monthly simulations of net ecosystem exchange and evapotranspiration from Orchidee, LPJ-GUESS, and RHESSyS (Biome-BGC is part of RHESSyS). LPJ, Orchidee, and Biome-BGC consistently predict higher GPP for the boreal forest by 10 to 23 %, lower GPP for temperate deciduous broadleaf forest and Mediterranean sites by 15 to 31% and 21 to 45% respectively. Between the models, LPJ is closest regarding the boreal forests (RMSE of 24%), Orchidee for temperate sites (RMSE of 16 and 27 % for conifers and broadleaves respectively), and Biome-BGC for Mediterranean evergreens (RMSE of 21 and 28 % for conifers and broadleaves respectively). The latter statement is somewhat ambiguous, given the small number of data points in the Mediterranean.

Forest ecosystem type	Number of sites	Mean GPP [gC/m2/yr]				Relative RMSE [%]		
		Observed	LPJ	Orchidee	Biome-BGC	LPJ	Orchidee	Biome-BGC
All	**37**	**1400**	**1097**	**1252**	**1243**	**32.34**	**29.56**	**29.65**
Boreal evergreen needleleaf	9	1003	1102	1225	1232	23.65	33.80	31.95
Temperate evergreen needleleaf	10	1643	1311	1537	1600	25.12	16.43	21.08
Temperate deciduous broadleaf	10	1534	1060	1305	1067	33.35	27.41	33.75
Mediterranean evergreen needleleaf	2	1586	879	894	1259	44.61	43.65	21.08
Mediterranean deciduous broadleaf	2	1197	811	558	665	42.35	60.84	51.32
Mediterranean evergreen broadleaf forest	4	1358	893	989	1097	41.03	32.59	28.29

Table 4-1: Relative RMSE and mean eddy covariance flux separated and modelled GPP, stratified by forest ecosystem type. The relative RMSE is calculated as RMSE divided by the mean of the eddy covariance flux separated GPP values. The model with smallest RMSE is underlined for individual forest types.

Declining GPP towards the Mediterranean region is primarily related to increasing dryness. Reichstein et al., 2007b found that GPP of forest ecosystems south of 52° latitude in Europe scales approximately linear with an index of water availability (IWA) which is defined as the ratio of actual to potential evapotranspiration. We find no systematic pattern of changes of the difference between simulated and eddy covariance based GPP along the gradient of water availability for this region, except that all three models tend to underestimate GPP (Figure 4-3). Underestimation of GPP in the water limited part of Europe suggests that the models do not simulate the soil moisture conditions appropriately (e.g. due to overestimation of evaporation and or transpiration) or are too sensitive to variations of soil moisture. However, we cannot rule out the effect of uncertain model input data. The Mediterranean is a very heterogeneous landscape and moisture conditions resulting from localised rainfall and local soil characteristics may deviate substantially from the rather coarse driver data. There is further indication that the meteorological data from REMO are biased towards too dry conditions. The REMO data show on average larger vapour pressure deficit and lower precipitation in southern Europe in comparison to an alternative meteorological dataset from ECMWF, which impacts strongly on simulations of GPP from Biome-BGC (Jung et al., in review).

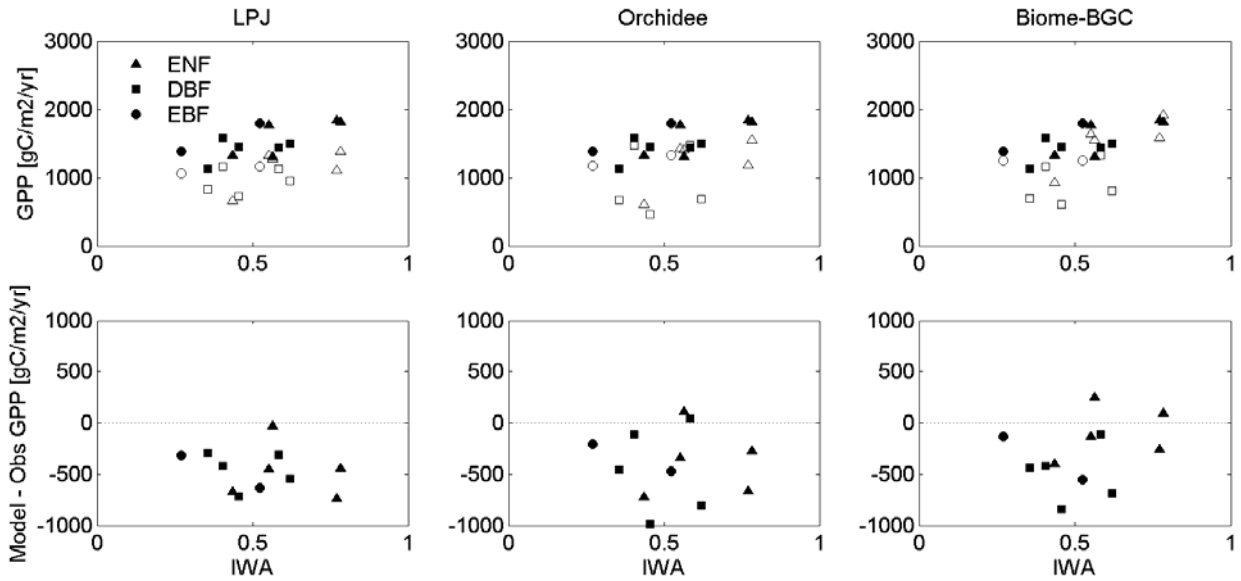


Figure 4-3: Top panel: eddy covariance flux separated (filled markers) and modelled (open markers) GPP along a gradient of water availability for sites south of 52° latitude. The index of water availability (IWA) is calculated as the ratio of actual to potential evapotranspiration and is based on measurements from the flux towers (see Reichstein et al., 2007b for details). Bottom panel: difference between modelled and eddy covariance flux separated GPP along the gradient of water availability. ENF: evergreen needleleaf forests, DBF: deciduous broadleaf forests, EBF: evergreen broadleaf forests.

Consistent with results of Morales et al., 2005, the discrepancy between simulations and reference data is higher for deciduous than for evergreen forests. The model's capacity to simulate the phenology of deciduous trees is therefore a likely factor that causes larger deviations for deciduous forests. Phenology involves several aspects relevant to carbon assimilation. The timing of budburst and leaf senescence determines the length of the growing season and together with the seasonal course of fAPAR the amount of light that can be harvested. Depending on the meteorological conditions, the timing of the onset of photosynthesis and thus transpiration may further impact on the efficiency to assimilate carbon later in the season due to the continuing depletion of available soil water. Beyond the seasonal course of LAI there is an internal 'physiological' phenology of leaf properties such as leaf nitrogen concentration and chlorophyll content that control photosynthetic capacity. Orchidee is the only model among the three considered in this study that accounts for such indirect effects using a dependence of maximum photosynthetic capacity on leaf age, which may explain why Orchidee performs better for temperate deciduous forests. A systematic test of the model's ability to simulate effects of phenology on photosynthesis at different sites as well as separating the relevance of different factors involved is challenging but needed for the future. Such study would require substantially more information of the forest ecosystem, including daily measurements of light absorbtion in the

canopy, and model simulations with daily output that are forced by in-situ measured meteorological and soil data.

Our primary goal is to assess the general correspondence of European scale simulations and eddy covariance based GPP along the MAT gradient. Thus we used the same driver data as previous modelling studies of Carboeurope-IP (Jung et al., in review; Trusilova et al., in review). This approach has the advantage that model evaluation is facilitated at their scale of application, i.e. continental to global including all uncertainties involved in large scale modelling. However, it trades-off to some extent with the identification of model structural uncertainties and unambiguous identification of which model performs best since input data effects can not be separated. Substantial deviation between the rather coarse soil and meteo input data and in situ conditions at the measurement sites can be expected due to small scale variability (esp. convective rainfall, cloudiness, soil structure and depth) and general uncertainties regarding the quality of the coarse scale model input. Considering input data effects and uncertainties of the GPP estimates from Carboeurope sites, the absolute simulated GPP values may be considered to be in the range of the uncertainty of our approach. Complementary, to this extensive data-model comparison study that covers well large climate gradients of Europe, we are currently undertaking effort to better understand real and model world controls of GPP variations for a few selected sites using in-situ measured model driver data.

4.3.2 Leaf Area Index

In this section we compare simulated maximum LAI with measurements in order to gain more insight in the model performances and what may cause some of the consistent discrepancy between eddy covariance based and modelled GPP particularly along the gradient from boreal to temperate climate.

LPJ and Orchidee simulate hardly any changes of LAI (expressed as fAPAR, see section 4.2.1) from the boreal to the temperate zone which results in substantial overestimation of fAPAR in the boreal zone but reasonable agreement for temperate forests (Figure 4-4). Biome-BGC captures the pattern qualitatively and does simulate an increase of LAI from boreal to temperate but not as strong as suggested by the measurements. The simulated LAI of boreal conifers is still too high

while LAI of temperate conifers appears too low. In addition, deciduous forests exhibit far too low leaf area in Biome-BGC. The measurements and all three models suggest decreasing LAI when moving from temperate to Mediterranean climate.

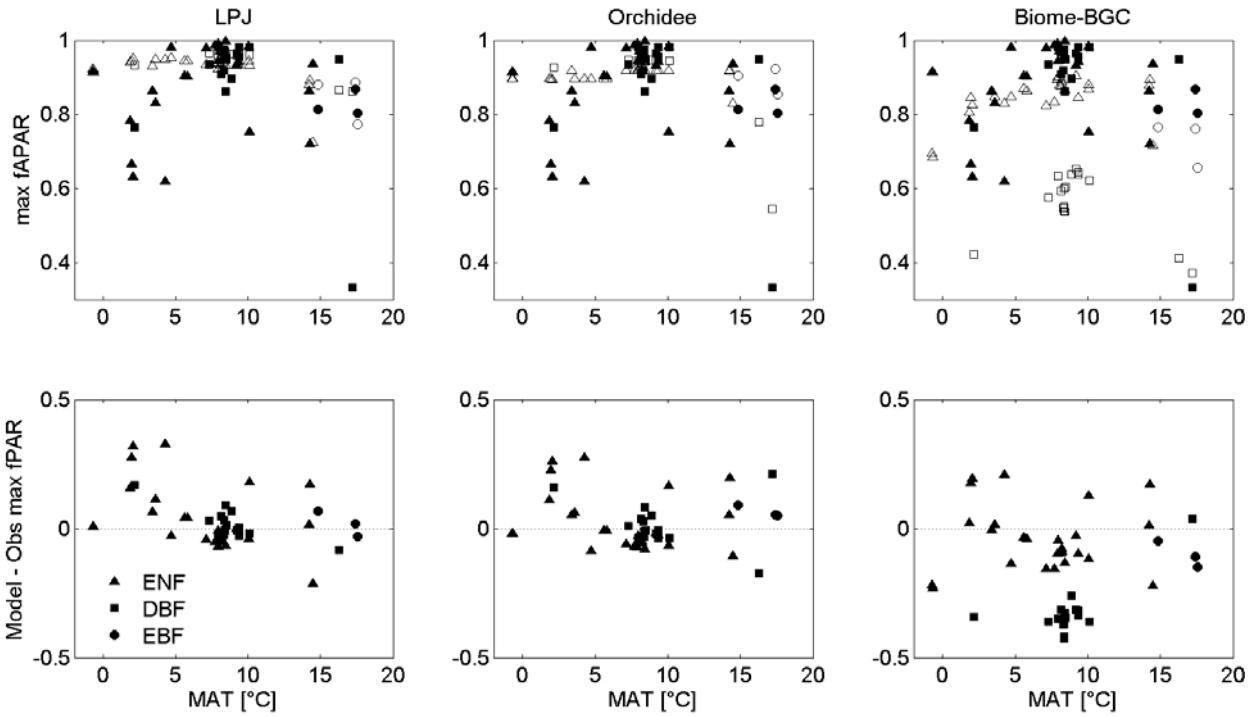


Figure 4-4: Top panel: observed (filled markers) and modelled (open markers) maximum fAPAR along the mean annual temperature gradient across Europe. Bottom panel: difference between modelled and observed fAPAR along MAT. ENF: evergreen needlleaf forests, DBF: deciduous broadleaf forests, EBF: evergreen broadleaf forests.

Leaf area is constrained by the availability of resources (Cowling and Field, 2003). In LPJ and Orchidee, the main resource limitation is plant available water while Biome-BGC includes nitrogen limitation. In a global NPP model intercomparison, Bondeau et al., 1999 suggested that models that include only water limitation tend to overestimate light harvesting when nitrogen limitation is present. The boreal zone is known to be nitrogen limited and this limitation decreases as nitrogen availability increases towards the temperate zone due to higher turnover but also anthropogenic deposition. The lack of an explicit nitrogen cycle may cause that LPJ and Orchidee do not simulate increasing LAI from boreal to temperate. On the other hand, the observed increase of LAI from boreal to temperate is partly an effect of a change in the prevailing conifer species from pine to spruce the latter being known to exhibit very high LAI (e.g. Breda, 2003) while global models cannot account for such species related effects. In the following

section we investigate to what extent the overestimation of LAI for the boreal forests may be responsible for the overestimation of GPP.

4.3.3 Decomposing GPP into APAR and RUE

Figure 4-5 shows APAR and RUE along MAT for boreal and temperate conifers. Because the simulated seasonal pattern of LAI was used to estimate site APAR, a site APAR for each model is presented (see section 4.2.3). Site and modelled APAR increase with MAT and are correlated significantly (Pearson's correlation, $p < 0.05$), but the site APARs increase more steeply with MAT (see also Table 4-2). As shown above, the models cannot reproduce the increase of fAPAR (i.e. increase of LAI) from boreal to temperate so that their slope of APAR vs MAT simply represents increasing radiation, while the larger observed slope is due to additionally increasing fAPAR.

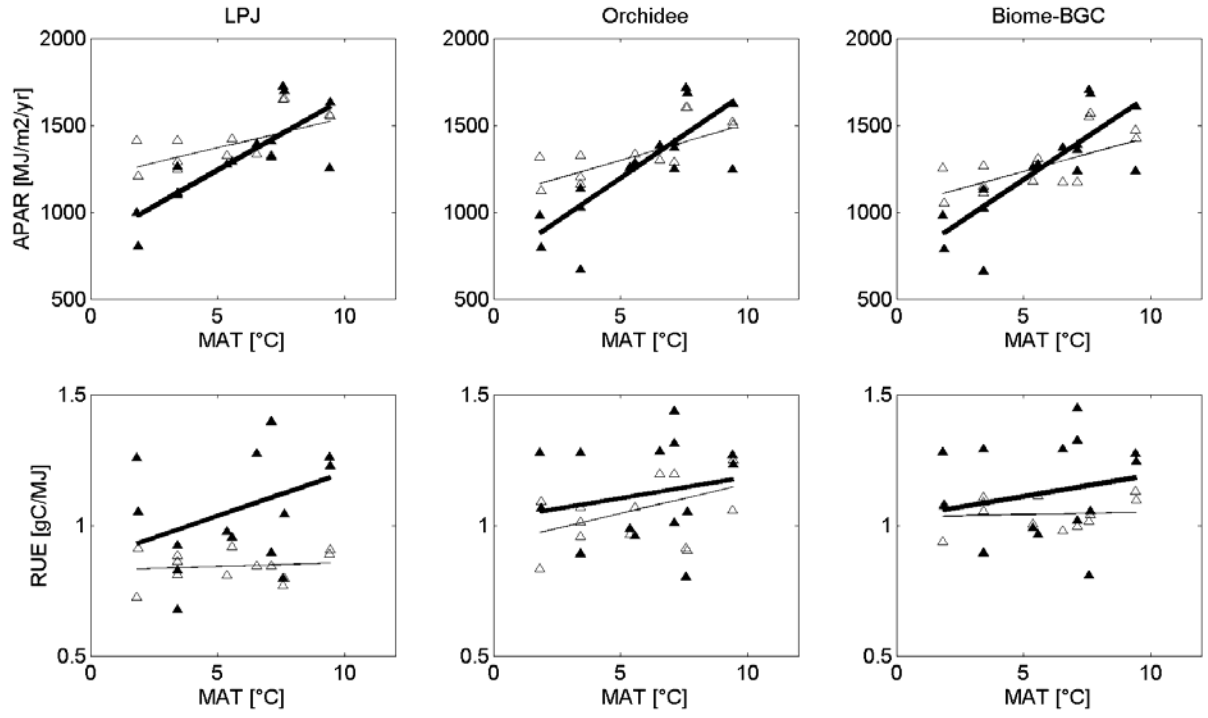


Figure 4-5: Site (filled markers) and modelled (open markers) trends of APAR and RUE along the mean annual temperature gradient for boreal and temperate coniferous forests. Bold line: trend of site values; thin line: trend of modelled values. The trend of site values can differ among models since the model specific simulated seasonal pattern of LAI was used to estimate APAR and consequently RUE.

Despite considerable scatter there is a trend of site RUE to increase with MAT which is only reproduced by the Orchidee model but the Pearson's correlation between MAT and RUE is not

significant in both cases. The trend of increasing RUE in observed data from boreal to temperate regions is confirmed by an independent study using remotely sensed fAPAR and in-situ measured radiation (Jung, unpublished). Rising RUE may result from more favourable temperature conditions for photosynthesis or due to increasing rubisco concentrations in the needles as nitrogen becomes more available. The latter factor is supported by data from Wright et al., 2004b that show larger concentrations of nitrogen per unit of leaf area in temperate than in boreal biomes. The Orchidee model shows increasing RUE with MAT likely because different optimum temperatures are assigned for boreal and temperate coniferous trees.

	Slope APAR vs MAT [MJ/°C]		Slope RUE vs MAT [gC/MJ/°C]	
	Modelled	Estimated from Observations	Modelled	Estimated from Observations
LPJ	34.58	82.78	0.003	0.033
Orchidee	42.11	99.65	0.016	0.023
Biome-BGC	40.1	97.87	0.002	0.017

Table 4-2: Trends of APAR and RUE along MAT for boreal and temperate evergreen needleleaf forests.

Site APAR and RUE for LPJ are different than 'site' for Orchidee and Biome-BGC, the latter two being almost identical (Figure 4-5). This difference can only result from different seasonal patterns of LAI. The assumption in LPJ that leaf area is constant over the year for evergreens seems to have a significant effect. Modelling small increases of fAPAR during summer (fresh needles) when radiation is high seems to be important for the magnitude of absorbed radiation.

We showed above that both site APAR and RUE increase more strongly with MAT than predicted by the models. The question is which of the two factors has the larger effect in explaining increasing GPP from boreal to temperate forests. Since GPP is the product of APAR and RUE, the answer to the question can be inferred from the coefficient of variation (standard deviation divided by mean) of both factors. The factor that varies more also controls more the variations of GPP. Site data and the models agree that changes of APAR is the dominant factor that explains increasing GPP from boreal to temperate coniferous forests in Europe, while changes of RUE are of secondary importance (Figure 4-6). The variation of APAR is more than twice as high as the variation of RUE and it is therefore likely that the data-model mismatch for boreal conifer forests is primarily caused by overestimating LAI. Since both, foliage area as well

as RUE is related to nitrogen availability the implementation or improvement of a nitrogen cycle in the models would likely enhance the model's performance. In fact, Magnani et al., 2007 have shown that the relationship of forest GPP along mean annual temperature in Europe is concomitant on nitrogen availability.

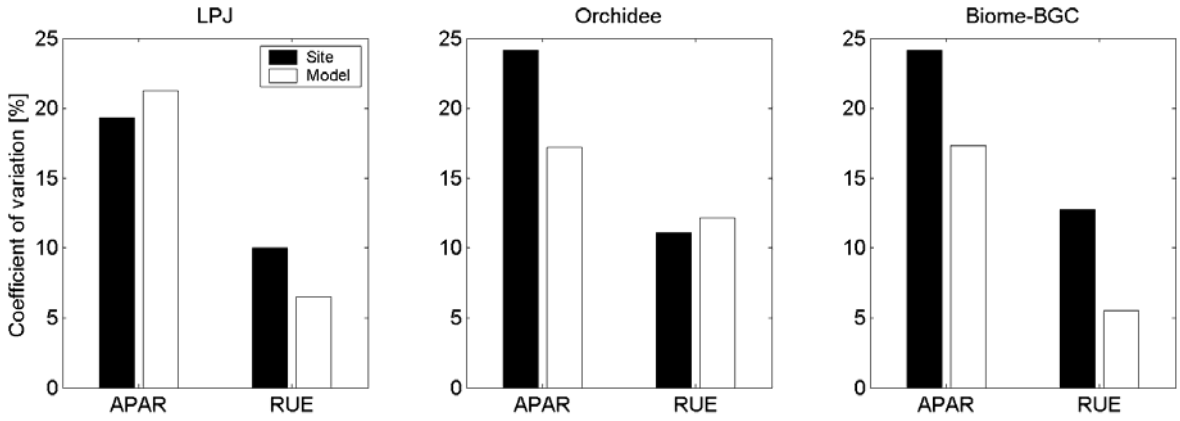


Figure 4-6: Coefficient of variation (standard deviation/mean) of APAR and RUE for boreal and temperate coniferous forests based on site and modelled data. The discrepancy of LPJ site data with site data based on Orchidee and Biome-BGC results from the assumption of constant leaf area over the year (see text).

Bias of GPP simulations along large environmental gradients is likely also related to assumptions made by representing vegetation using broad categories of plant functional types (PFTs). Many important plant traits (e.g. leaf nitrogen concentration, specific leaf area, leaf longevity) that control biogeochemical cycling are represented as constant PFT specific parameters in the models. These traits are known to vary within and between PFTs, and systematically along environmental gradients (e.g. Reich and Oleksyn, 2004; Wright et al., 2005; Wright et al., 2004b; Wright et al., 2006). Accounting for the variation of vegetation properties which are currently kept constant in the models would certainly improve their predictability. Using simple empirical relationships with climate have not improved simulations successfully (e.g. White et al., 2000). However, approaches of understanding the variation and co-variation of key plant traits using the theory of optimality in ecosystems regarding the use of resources (mainly water, light, nitrogen) has been promising (e.g. Anten, 2002; Anten, 2005; Hikosaka, 2005; Shipley et al., 2006). This concept is attractive for global prognostic ecosystem models but there is still too little known regarding when optimality applies, what is optimised and how, and the respective time scale.

4.4 Conclusions

We estimate the root mean square error of prediction (RMSE) over all forest sites to be in the order of 420 gC/m2/yr (~ 30%) for all three models. In terms of absolute simulated GPP values this uncertainty range may be considered to be within the joint uncertainty resulting from input driver data and eddy-covariance based GPP estimates. However, we find systematic biases in the model simulations along the climatic gradient from the boreal to the Mediterranean region.

Based on a simple method that decomposes GPP into APAR and RUE, we conclude that the tested models consistently overestimate GPP for boreal forests due to the tendency of the models to simulating too high LAI in this region. Due to general N-limitation in the boreal zone, accounting explicitly for nitrogen limitation should reduce the simulated LAI and therefore improve the model performance for the boreal zone. The method of GPP decomposition may be useful for future evaluations of large scale carbon cycle simulations based on global measurement databases of GPP that include also LAI data.

The tendency of all three models to underestimate GPP in the water limited part of Europe indicates issues of model structure regarding their soil hydrology. However, this pattern is likely, at least partly, a consequence of questionable meteorological input data over this region.

We have undertaken an evaluation of global ecosystem models on a continental scale, including many sites and covering large climatic gradients. Such effort has been neglected in the past but is necessary to identify model biases along environmental gradients or to gain confidence in simulations. Large scale data-model comparison studies need to be fostered by the community in the future.

5 Diagnostic assessment of European gross primary production[4]

Abstract

We present an approach to estimate gross primary production (GPP) using a remotely sensed biophysical vegetation product (**F**raction of **A**bsorbed **P**hotosynthetically **A**ctive **R**adiation, FAPAR) from the European Commission Joint Research Centre (JRC) in conjunction with GPP estimates from eddy covariance measurement towers in Europe. By analysing the relationship between the cumulative growing season FAPAR and annual GPP by vegetation type we find that the former can be used to accurately predict the latter. The root mean square error of prediction is of the order of 250 $gC/m^2/yr$, which is lower than reported errors of existing GPP models. The cumulative growing season FAPAR integrates over a number of effects relevant for gross primary production such as the length of the growing season, the vegetation's response to environmental conditions, and the amount of light harvested which is available for photosynthesis.

Corroboration of the proposed GPP estimate (noted **F**APAR based **P**roductivity **A**ssessment + **L**and **C**over, FPA+LC) on the continental scale with results from the MOD17+ radiation use efficiency model, an artificial neural network up-scaling approach (ANN), and the **L**und-**P**otsdam-**J**ena **m**anaged **L**and biosphere model (LPJmL) supports our suggested approach. The spatial pattern of mean annual GPP compares favourably with the estimates from ANN (R^2=0.74). Total GPP over the European model domain as estimated by the four different models ranges from 7.07 to 8.72 PgC/yr or within ~20%. Accounting for bias resulting from the meteorological input data used to drive MOD17+, ANN, and LPJmL the four models converge to a total GPP of the European domain of 8.29 to 8.72 PgC/yr, i.e. they fall within ~5% of each other. While our analysis suggests that results from data-driven models may be used to evaluate process-driven models regarding the mean spatial pattern of GPP, there is too little consensus among the diagnostic models for such purpose regarding interannual variability.

[4] Manuscript in review: Jung, M., Verstraete, M., Gobron, N., Reichstein, M., Papale, D., Bondeau, A., Robustelli, M., Pinty, B.: Diagnostic Assessment of European Gross Primary Production. *Global Change Biology*, in review.

A major advantage of the FPA+LC approach presented in this paper is that it requires no additional meteorological input driver data that commonly introduce substantial uncertainty. The FPA+LC GPP product is attractive for various applications such as evaluating biosphere models on the continental scale, and quantification of GPP over large regions. The FAPAR-based GPP product is available upon request from the first author.

5.1 Introduction

Gross primary productivity (GPP) is the flux of carbon into ecosystems via photosynthetic assimilation. Respiratory fluxes and distribution of carbon to different compartments depend on this initial quantity entering the system. Recent studies have highlighted the significance of GPP in driving the net carbon balance, both in terms of spatial as well as temporal variations (Ciais et al., 2005; Luyssaert et al., 2007; Reichstein et al., 2007b; van Dijk et al., 2005). GPP is thus a critical flux that drives the carbon budget of ecosystems.

At the local scale, methods have been developed that allow one to estimate GPP on the basis of eddy covariance measurements of net ecosystem exchange (NEE) by separation into the gross fluxes GPP and ecosystem respiration (Reichstein et al., 2005, Desai et al., 2007). Such estimates of GPP are limited to distribution of FLUXNET sites (http://www.fluxnet.ornl.gov). Quantifying and studying GPP variations at continental to global scale requires some sort of spatially explicit modelling. In principle, two groups of such models can be distinguished: (1) diagnostic or data-driven approaches, and (2) prognostic or process-oriented ecosystem models.

Prognostic models simulate the carbon cycle of ecosystems based on mechanistic and semi-empirical process formulations. Although the process of photosynthesis is well understood at the leaf scale, process-oriented models deviate substantially in their simulations of GPP both spatially and inter-annually (Jung et al., 2007c). Studying and reducing the uncertainties of prognostic models is crucial to gain confidence of predictions of the evolution of the Earth system including simulated carbon cycle climate feedbacks (Friedlingstein et al., 2006b). Evaluating process-based biosphere models requires readily available accurate data on the scale of their application, i.e., continental to global, which is still lacking. Results of diagnostic models

may close this gap if they can provide more accurate GPP estimates. This, however, has not been exemplified yet.

Diagnostic models use spatial fields of remotely sensed vegetation properties to scale-up local estimates of GPP. Radiation use efficiency models (Monsi and Saeki, 1953; Monteith, 1965; Running et al., 2004; Xiao et al., 2004) are most commonly used where GPP is estimated as the product of absorbed photosynthetic active radiation (APAR) and radiation use efficiency (RUE). RUE is usually calculated as a land cover-specific property that is reduced by scalars according to meteorological or soil hydrological conditions. The estimation of RUE constitutes the largest conceptual uncertainty of radiation use efficiency models. Recent efforts to assess RUE directly from space data by means of fluorescence or the photochemical reflectance index have just begun (see thorough review by Grace et al., 2007). A number of issues still preclude an operational use on the continental scale. Up-scaling carbon fluxes from FLUXNET sites to the continent by means of remotely sensed vegetation properties and meteorological data using artificial intelligence have been proposed (Papale and Valentini, 2003; Yang et al., 2007). Recently, Beer et al., 2007 have introduced a method to estimate GPP of watersheds based on its water balance. The authors use empirical relationships of the ecosystems water use efficiency to Leaf Area Index (LAI) and soil water holding capacity that are applied spatially using remotely sensed LAI and gridded soil data. In conjunction with rainfall, vapour pressure deficit, and river runoff data, GPP can then be approximated.

Prognostic and diagnostic models can generally be tuned to accurately reproduce GPP at the site level; a major obstacle for the modelling over large scales is the generalisation of parameter sets for vegetation types. Moreover, prognostic and diagnostic models are sensitive to uncertainties in input data, especially meteorological forcing fields. The choice of the meteorological input data set alone can result in a 20% difference of simulated GPP as estimated by Jung et al., 2007c for Europe using the Biome-BGC model (Thornton, 1998) and the globe by Zhao et al., 2006 with the MOD17 model (Running et al., 2004). However, structural uncertainties of different modelling approaches are found to even exceed the effect of different input data choices (Jung et al., 2007c). Estimating GPP over large regions remains a challenge and confidence can only be gained by comparing different approaches and studying their individual uncertainties.

Given the uncertainties of existing modelling approaches to predict GPP over large regions accurately, there is renewed interest to directly relate remotely sensed vegetation properties to GPP. Relationships between integrated Normalized Difference Vegetation Index (NDVI) over the growing season and net primary productivity (NPP) had already been reported in the 1980s for regions in North America (Box et al., 1989; Cook et al., 1989; Goward et al., 1985). Sims et al., 2006 suggested that the Enhanced Vegetation Index (EVI) from Moderate Resolution Imaging Spectroradiometer (MODIS) is a better predictor for daily GPP than the MOD17 GPP product if only growing season data points are compared.

Remote sensing-based vegetation products have an obvious potential for ecosystem productivity prediction on continental to global scale (e.g. Cao et al., 2004; Goetz et al., 2000; Hicke et al., 2002; Knorr and Heimann, 1995; Potter et al., 1993; Prince and Goward, 1995; Ruimy et al., 1996; Running, 1994; Xiao et al., 2004). Advances in satellite sensor technology and physically-based radiation transfer algorithms now allow a much improved retrieval of biophysical vegetation properties (like the Fraction of Absorbed Photosynthetically Active Radiation, or FAPAR) that can be evaluated against ground measurements. This advanced remote sensing based vegetation product, used in conjunction with networks of eddy covariance flux measurement sites and standardized data processing chains (Papale et al., 2006), offers unprecedented possibilities to investigate and exploit relationships between remotely sensed vegetation properties and gross carbon uptake of ecosystems at the continental scale.

In the first part of this study we develop an empirical model to estimate annual sums of GPP over Europe based on remotely sensed FAPAR and eddy covariance flux tower measurements. In the second part we apply this model to the European domain and corroborate our results with independent simulations from the LPJmL biosphere model, the radiation use efficiency model MOD17+, and a neural network based up-scaling of GPP. To our knowledge, this is the first time that results from different data-driven GPP models are compared at the continental scale, which allows us (1) to evaluate to what extent diagnostic models can be used as a reference for process-oriented models, (2) to provide a realistic bound of European GPP.

5.2 Relating the cumulative growing season FAPAR to gross carbon uptake

5.2.1 Materials and Methods

GPP estimates from eddy covariance flux tower measurements

GPP is estimated by separating the measured net flux of carbon between the land surface and the atmosphere (net ecosystem exchange, NEE) into its gross constituent fluxes GPP and terrestrial ecosystem respiration (TER).

NEE = TER - GPP Equation (5-1)

The flux separation follows Reichstein et al., 2005 where night-time temperature sensitivities are determined within short-term periods and extrapolated to the daylight period. This allows for the quantification of ecosystem respiration. GPP is then given by the difference between ecosystem respiration and net ecosystem exchange.

Annual sums of GPP based on flux separated eddy covariance measurements of NEE are subject to various uncertainties that may be introduced by a number of processing steps: u*-filtering, spike removal, storage correction (Papale et al., 2006), gap-filling (Moffat et al., 2007), partitioning into of NEE into GPP and TER (Desai et al., 2007; Papale et al., 2006). Effects of problematic micrometeorological conditions that are not filtered out by the u*-thresholds remain under intense study and can introduce considerable errors, but seem to be confined to specific site conditions (Aubinet et al., 2005; Marcolla et al., 2005).

Joint uncertainties are surely site-specific but are usually within 100 $gC/m^2/yr$. We follow Reichstein et al., 2007b who use an uncertainty of $200gC/m^2/yr$ for annual GPP sums as a conservative estimate.

JRC-FAPAR from the SeaWiFS sensor

FAPAR is the fraction of absorbed radiation in the PAR domain by green vegetation. Following the conceptual work of Pinty et al., 1993 and Verstraete and Pinty, 1996 on optimized vegetation indices, Gobron et al., 2000 proposed a generic scheme to produce equivalent, and thus

comparable, FAPAR products derived from various optical sensors to achieve long time series of FAPAR products from space instruments. The JRC algorithm capitalizes on the physics of remote sensing measurements and minimizes contaminating effects of sun-target-sensor geometry, atmospheric aerosol, and soil brightness changes. Basically, the useful information on the presence and state of vegetation is derived from the red and the near-infrared spectral band measurements. The information contained in the blue spectral band, which is very sensitive to aerosol load, is ingested in order to account for atmospheric effects on these measurements. In practice, the generic FAPAR algorithm implements a two step procedure where the spectral BRFs measured in the red and near-infrared bands are, first, rectified in order to ensure their optimal decontamination from atmospheric and angular effects and, second, combined together to estimate the FAPAR value. The protocol for the validation of SeaWiFS FAPAR products has been proposed in Gobron et al., 2006 and the results show that the accuracy is at about ±0.1, when comparing against ground-based estimates. Additional analyses, achieved with the MERIS instrument, shows that the impact of the top-of-atmosphere radiance uncertainties on the products is less than 10% (Gobron et al., 2007).

The SeaWiFS-based JRC-FAPAR product currently covers the period from September 1997 to June 2006, with a nominal spatial resolution of 2 km and a temporal resolution of 10 days. This product is available from http://fapar.jrc.ec.europa.eu/. Time series of the JRC-FAPAR at the locations of CarboEurope sites are available as 10 day composites, either for the exact pixel or for a 3×3 pixel window. We do not use data from 2005 onwards because the original sensor radiances are not longer available at the full spatial resolution. We use extracts for the 3×3 windows since they provide a less noisy and more complete record, however this may dilute somewhat the relation between tower and satellite data. We account for this by excluding sites where the 1×1 and 3×3 FAPAR time series are very different from each other. We calculate Pearson's correlation coefficient and the 'modelling efficiency' measure (Tedeschi, 2006) and keep only sites with values larger than 0.7 and 0.5 respectively. We thus exclude sites where large local heterogeneity is anticipated and where the tower may not be representative for the larger area.

This processing yields a set of GPP data and FAPAR time series for 39 sites (117 site-years) (Figure 5-1). These 39 sites span various vegetation types as well as a large environmental gradient, from boreal to Mediterranean climates.

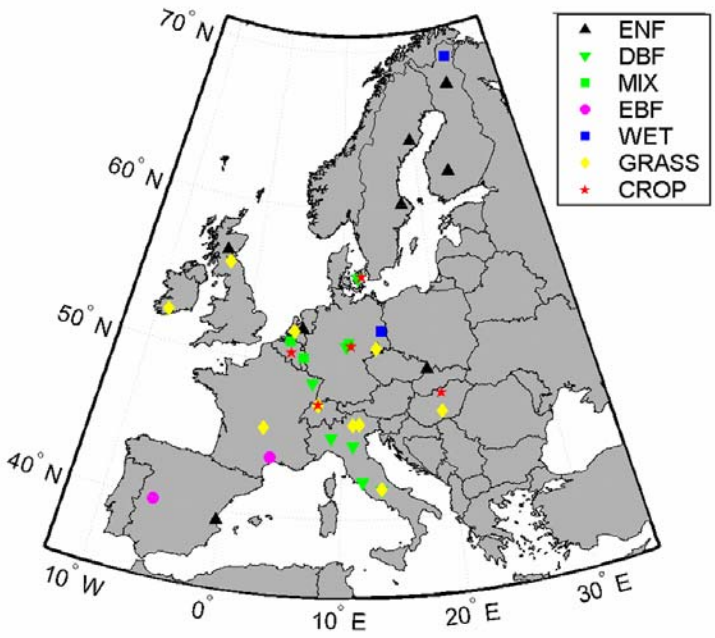


Figure 5-1: Map of Europe with CarboEurope sites used in this study.

Algorithm to calculate the cumulative FAPAR of the growing season

We use a simple gap filling approach where short gaps of maximum 3 consecutive FAPAR data points are replaced by linear interpolation. Long gaps of maximum 10 consecutive dates are replaced by the mean seasonal cycle when possible. Long gaps are commonly restricted to periods of snow cover or during polar night at high latitudes when the vegetation is dormant. Thus, uncertainties due to the filling of long gaps affect only rarely the calculated cumulative growing season FAPAR value.

The challenge is to extract the integrated FAPAR of the growing season from the FAPAR time series, since this constitutes the information that is sensitive to vegetation productivity. Several methods have been proposed to identify start and end events of the growing season from multi-temporal satellite data (e.g. Bradley et al., 2007; Duchemin et al., 1999; Sakamoto et al., 2005; Verstraete et al., 2007; White and Nemani, 2006; Zhang et al., 2003, see also review of Cleland et al., 2007). These algorithms are either based on thresholding the time series or on properties of

mathematical models describing the time series (e.g. inflection points). There are no explicit standards of how to define start and end events of a growing season. The choice of the criteria is a bit arbitrary and dependent on a particular application, also since the phenological behaviour is often not well described by 'events' but may be a rather fuzzy transition. Noise of space derived time series is a major challenge for a robust performance of a growing season length algorithm.

After extensive testing of various options to identify growing season start and end events we arrived at a simple but robust method that calculates the cumulative growing season FAPAR value without determining the start and end events of a growing season explicitly. Firstly, we sum the FAPAR values that are above a background value that is typical for non-growing season conditions (Figure 5-2). FAPAR usually does not decrease to zero because some PAR absorption of the land surface remains, during the dormant period. This 'background' value tends to vary among sites but is rather consistent among years at one site. We analyzed the relationship of the sum of FAPAR values above this background value with annual GPP and found that adding back the background values improves the predictability. In order to add back the background values we need to estimate the length of the growing season. We infer the length of the growing season from a geometrical solution using the accumulated FAPAR and the annual maximum FAPAR by assuming that the FAPAR record is shaped like half of an ellipse (Equation 5-2 and 5-3), i.e. similar to bell shaped which is a valid approximation in most cases. Given the area of the ellipse (twice the accumulated FAPAR), and the major axis of the ellipse (annual maximum FAPAR minus background), the minor axis of the ellipse can be calculated which equals half of the growing season length. The inferred length of the growing season is then used to add back the background value that has been initially subtracted (Equation 5-4, Figure 5-2). We estimate the uncertainty of the accumulated FAPAR value by summing the reported uncertainty of the FAPAR values of 0.1 (Gobron et al., 2006) over the growing season.

$$2 \times CUM_{BG} = \pi \times MAX_{BG} \times GSL/2 \qquad \text{(Equation 5-2)}$$

$$GSL = 4 \times CUM_{BG} / (\pi \times MAX_{BG}) \qquad \text{(Equation 5-3)}$$

$$CUM\ GSL\ FAPAR = CUM_{BG} + GSL \times BG \qquad \text{(Equation 5-4)}$$

where CUM_{BG} is the accumulated FAPAR of a year after the subtraction of the background value (BG), BG is estimated as the 10^{th} percentile of the gap filled FAPAR time series, MAX_{BG} is the maximum FAPAR value of a year minus the background, GSL is growing season length, CUM GSL FAPAR is the cumulative FAPAR of the growing season of a year.

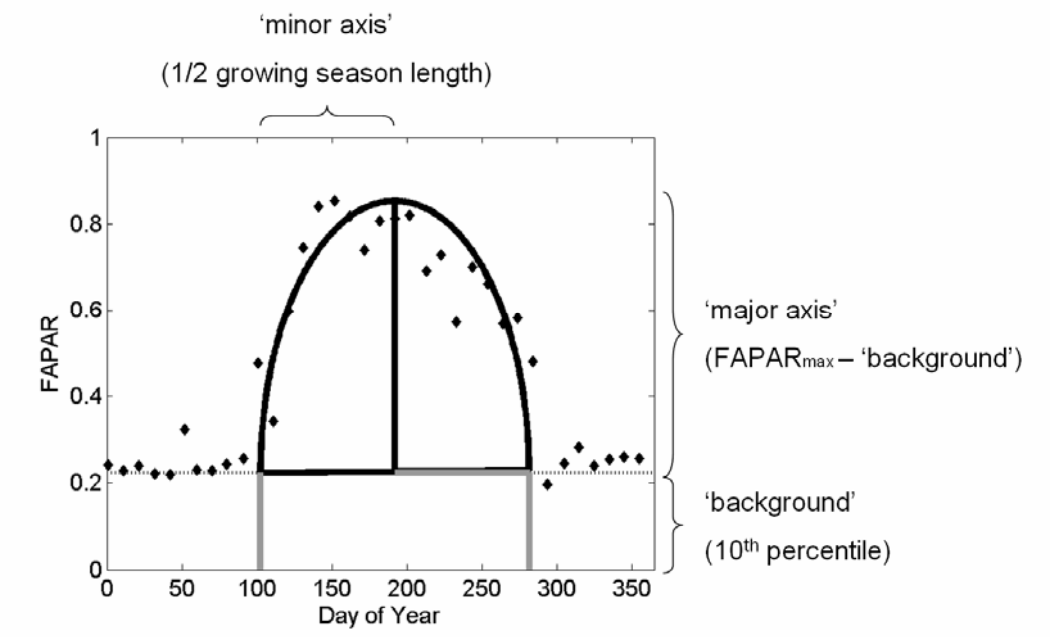


Figure 5-2: Illustration of the algorithm to calculate the cumulative FAPAR of the growing season. The cumulative FAPAR of the growing season is estimated as the sum of FAPAR values above the background (area of half of the ellipse) plus the length of the growing season times the background (area of the background rectangle). The length of the growing season is given by twice the minor axis of the ellipse.

Our method of inferring the length of the growing season is an approximation that may lead to imprecise results if the true shape deviates substantially from an ellipse or if multiple growing seasons are present within a year. However, it produces reliable patterns for Europe (Figure 5-3) and the uncertainty on the final cumulative growing season length FAPAR value is small since the bulk of the signal originates from the sum of FAPAR values larger than the background; the growing season length is only needed as an approximation to add back the background value. This simple method is computational efficiency and robust, even for noisy time series.

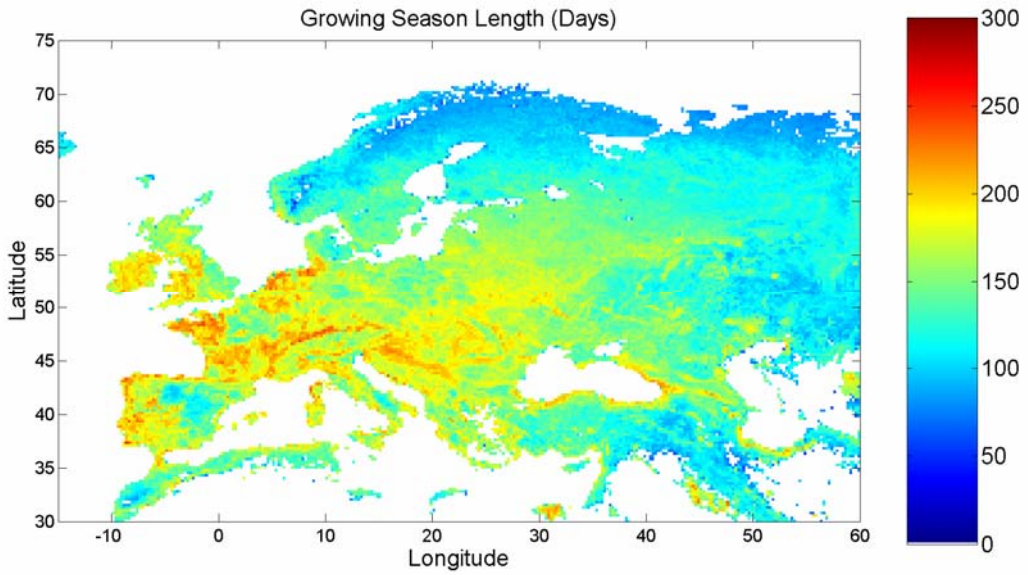


Figure 5-3: Map of mean growing season length (1998-2002) based on the proposed algorithm to calculate the cumulative FAPAR of the growing season.

5.2.2 Results and Discussion

The accumulated FAPAR of the growing season explains more than 50 % of the variance in annual GPP data (R^2 = 0.56, n=117) across different vegetation types and years (Figure 5-4a) when we fit a logarithmic function (Equation 5-5).

$$GPP = a \times ln\,(\text{CUM GSL FAPAR}) + b \qquad \text{(Equation 5-5)}$$

We investigated whether different and possibly stronger relationships of the same type exist within plant functional types and found that a stratification into herbaceous (wetlands, grasslands, and crops), evergreen forests (needle and broadleaf), mixed forests, and deciduous forests gave the best results (Figure 5-4b). The relationship becomes substantially stronger for herbaceous ecosystems (R^2=0.8) and evergreen forests (R^2=0.71). For mixed forests, the accumulated growing season FAPAR still explains more than 50 % of the variation of annual GPP while we find no such relationship for deciduous forests (Table 5-1).

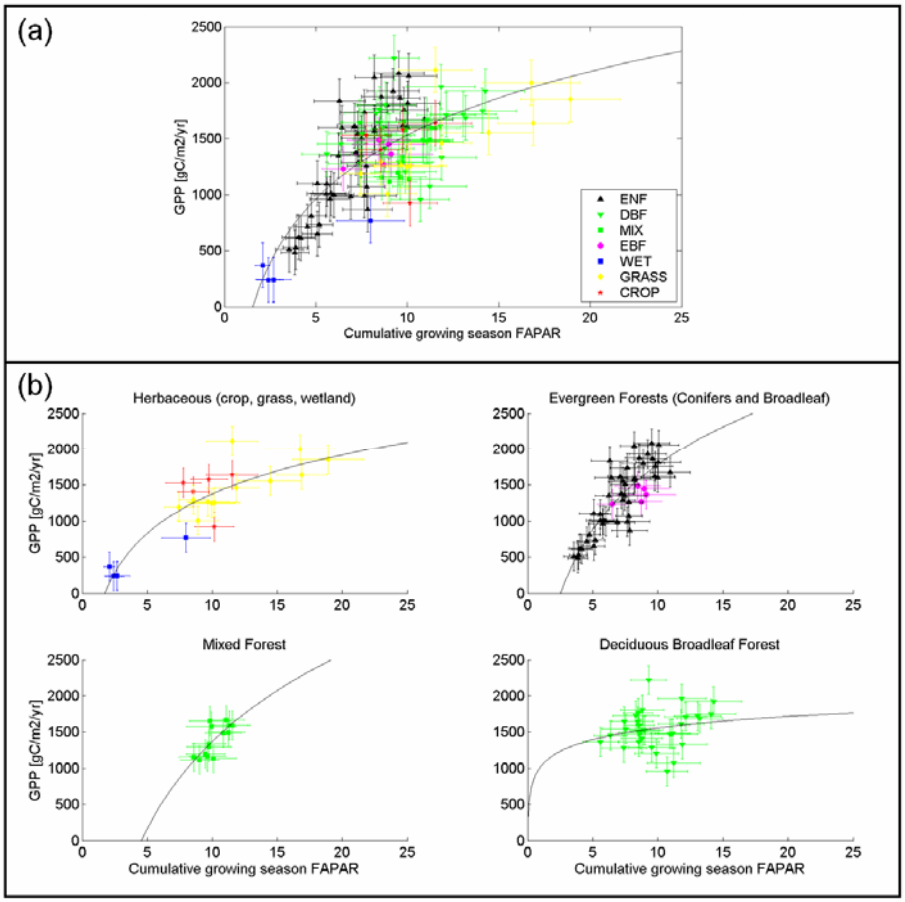


Figure 5-4: Scatter plots of the cumulative growing season FAPAR and GPP for (a) all data points, (b) stratified by ecosystem types. The curves correspond to the best fit.

	A	B	R^2	RMSE [gC/m^2/yr]	Relative RMSE	Number of site years	Number of sites
ALL	821.71	-360.02	0.56	280	0.20	117	39
Herbaceous	785.96	-434.66	0.80	242	0.20	22	17
Evergreen Forests	1301.8	-1211.1	0.71	243	0.19	49	11
Mixed Forest	1737.3	-2627.6	0.54	138	0.10	14	3
Deciduous Broadleaf Forest	230.96	1024	0.04	248	0.16	32	8

Table 5-1: Statistics on the relationship between the cumulative FAPAR of the growing season and annual sums of GPP for different groups of ecosystems. A and B are the parameters of the logarithmic fit from Equation 5-5. The relative RMSE is defined as the RMSE divided by mean GPP.

The root mean square errors (RMSE) are less than 250gC/m2/yr for these vegetation-specific functions. In comparison, the RMSE of three process-oriented ecosystem models (LPJ, Orchidee, Biome-BGC) to simulate between site variations of GPP of forest ecosystems in Europe has been quantified to be 414–453 gC/m2/yr (n=37, Jung et al., 2007a) and larger uncertainties are expected for herbaceous vegetation, in particular crops. The RMSE of the MOD17 GPP product

has been evaluated between 386 and 490 gC/m2/yr (R^2 between 0.56 to 0.74, Yang et al., 2007) and 388 – 414 gC/m2/yr (R^2 between 0.33 and 0.47) for the support vector machine approach of Yang et al., 2007 (Table 5-2). Reviewing the literature, we noticed that statistics of predictability are commonly reported for 8-daily values (temporal resolution of MODIS products) but not for annual sums of GPP. Interestingly, a good model performance for daily data does not necessarily translate into a good model performance for annual data, suggesting that consistent seasonal bias can play an important role for models using a daily time step (Table 5-2). The development of our regression model explicitly for the annual time scale is probably an important reason why our RMSEs are comparatively small.

We conclude that the relationship between the cumulative FAPAR of the growing season and GPP is a promising approach to scale up gross carbon uptake to large regions using the remotely sensed FAPAR data, without the need for additional meteorological input data. The uncertainty introduced due to the poor performance for deciduous forests is relatively small since deciduous broadleaf forests cover only 13 % of the European land surface; 80 % are covered by herbaceous vegetation and evergreen forests for which we can predict GPP accurately (vegetation areas calculated from the SYNMAP 1km land cover map (Jung et al., 2006)).

How are the cumulative FAPAR of the growing season and annual gross carbon uptake linked?
FAPAR is the fraction of absorbed photosynthetically active radiation absorbed by the green vegetation. Therefore, FAPAR is both a determinant of photosynthesis and consequence of vegetation's primary production. FAPAR controls carbon assimilation since it determines the amount of light that is available for carboxylation. A change of FAPAR scales linearly with a change of GPP for a given level of radiation and radiation use efficiency. Moreover, the seasonal course of FAPAR is concomitant on variations of 'greenness'. The greenness related vegetation index EVI was found to co-vary with radiation use efficiency and this co-variation was found important for the close correspondence of EVI and GPP (Nakaji et al., 2007; Sims et al., 2006). Therefore, the JRC-FAPAR should also account implicitly for some of the variation of RUE, especially seasonal changes.

Model name	Model input	Vegetation types	RMSE 8-daily [gC/m^2/d]	R^2 8-daily	RMSE annual [gC/m^2/yr]	R^2 annual	N data points annual	Reference
Artificial neural network	FAPAR (MODIS), Temperature (NCEP-REMO), Radiation (NCEP-REMO), VPD (NCEP-REMO)	Cropland	2.99	0.66	-	-	-	Papale, unpubl.
		Deciduous Broadleaf forest	2.00	0.80	-	-	-	
		Evergreen broadleaf forest	1.14	0.70	-	-	-	
		Evergreen needleleaf forest	1.71	0.74	-	-	-	
		Grass-/wetland	2.00	0.65	-	-	-	
MOD17 optimised with CarboEurope flux tower measurements	FAPAR (MODIS), Radiation (in-situ), Temperature (in-situ), VPD (in-situ)	Various	1.49	0.77	-	-	-	Reichstein, unpubl.
		Temperate Evergreen needleleaf forest	1.1	0.85	-	-	-	
		Deciduous Broadleaf forest	2.08	0.81	-	-	-	
		Mediterranean Evergreen broadleaf forest	1.29	0.69	-	-	-	
		Mediterranean Evergreen needleleaf forest	1.68	0.49	-	-	-	
		Mixed Forest	1.14	0.82	-	-	-	
		C3 cropland	1.38	0.78				
Support Vector Machine	LST (MODIS), EVI (MODIS) Radiation (in-situ)	various	1.87	0.71	388	0.47	15	Yang et. al. 2007
		Non-forest	2.05	0.63	369	0.44	9	
		Forest	1.63	0.79	414	0.33	6	
MOD17	FAPAR (MODIS), Radiation (DAO), Temperature (DAO), VPD (DAO)	various	-	-	451	0.56	15	
		Non-forest	2.71	0.39	490	0.58	9	
		Forest	2.08	0.67	386	0.7	6	
		various	-	-	-	0.74	22	Heinsch et al. 2006
		all	-	0.58*	-	-	-	Sims et al. 2006
		cold climate evergreen	-	0.66*	-	-	-	
		west coast evergreen	-	0.20*	-	-	-	
		deciduos forest	-	0.59*	-	-	-	
		grassland	-	0.33*	-	-	-	

Table 5-2: Compilation of RMSE and R^2 values for data-driven GPP models from multi-site studies using eddy covariance GPP estimates. The R2 values reported by Sims et al. refer to 16 day averages rather than 8 day averages and are labelled with a star. The RMSE and R^2 on annual scale from Yang et al. 2007 were calculated from Table 4 in Yang et al. 2007.

The functional convergence hypothesis (Field, 1991; Goetz and Prince, 1999) suggests that GPP scales linearly with APAR due optimisation of resource acquisition, allocation, and use in ecosystems that results in a narrow range of radiation use efficiencies. If we assume this hypothesis to be correct, a nearly linear relationship between the accumulated FAPAR and GPP exists, because FAPAR is the dominant control of APAR (i.e. how much light is harvested not how much light is incident). Indeed, absorbed radiation is strongly correlated with GPP for herbaceous vegetation and evergreen forests, while incident radiation is only weakly or not correlated with GPP (Table 5-3). However, the relationship between the cumulative growing season FAPAR and GPP is stronger than the relationship between absorbed radiation and GPP. The latter suggests that FAPAR is also consequence of GPP, which is outlined in the next paragraph.

FAPAR reflects the amount of photosynthetic active tissue of the land surface, which is directly dependent on (past) productivity. In this sense, FAPAR would most closely be related to foliage NPP which in turn is strongly correlated with GPP. Thus, FAPAR senses the vegetation's response to environmental conditions, and by integrating the FAPAR values over the growing season we account also for lagged effects that may occur, for instance after a period of water stress. Given that the cumulative FAPAR of the growing season accounts for a number of factors relevant to plant productivity, it can be expected to be an excellent indicator for GPP and NPP. The relationship of GPP with the cumulative FAPAR of the growing season is stronger than with climate related determinants of GPP in Europe (mean annual temperature, index of water availability, Reichstein et al., 2007b, Table 5-3). However, Reichstein et al., 2007b pointed out that annual GPP of forest ecosystems in Europe follows mean annual temperature at northern sites and water availability in southern sites. Such stratification in broad climatic zones is not necessary when the cumulative growing season FAPAR is used to predict GPP since the effect of climatic conditions are already integrated. Instead, stratification into vegetation types improves the prediction of GPP by the cumulative growing season FAPAR.

In general, it cannot be assumed that the CarboEurope flux tower measurements and space-derived measurements of FAPAR actually sample exactly the same vegetation. Flux towers have a typical foot print radius of a few hundreds of meters for forest sites and a few tens of meters for shorter herbaceous sites, while we use a 6×6 km (3×3 pixels) sampling for the FAPAR data. The

strong relationships between the remotely sensed cumulative FAPAR of the growing season and flux tower based GPP suggests that the local productivity as seen by the tower varies in concert with its surrounding, which is also observed by the satellite. This can be expected since environmental factors and resources (meteorological, soil, and nutrient conditions) usually do not vary too much between the two scales, at least in the context of continental scale environmental gradients. Moreover, we excluded sites where large differences of the 1×1 and 3×3 pixel FAPAR time series exist, which indicated large local landscape heterogeneity. That the satellite samples a larger area around the tower may in fact contribute to the strength of the relationship between the cumulative FAPAR and GPP for forest sites. The herbaceous fraction within the FAPAR samples for forest sites (under storey, surrounding crop/grassland) may act as bio-indicators for the productivity of the forest and its seasonality may enhance the growing season signal at evergreen forests (cf. Sims et al., 2006). The FAPAR signal of herbaceous vegetation is more sensitive to the variability of environmental conditions since herbaceous plants respond fast to e.g. water stress by yellowing or senescence. Trees in contrast experience similar stress which results in reduced photosynthesis but do not necessarily react with leaf yellowing or shedding that the FAPAR would pick up. Also Reichstein et al., 2007a attributed FAPAR changes in evergreen needle-leaf forests during the 2003 summer heat wave largely to leaf yellowing of herbaceous plants (understorey, mixed pixels). In summary, the cumulative FAPAR of the growing season can be 1) cause, 2) effect, or 3) indirect indicator of GPP. Clearly, we cannot separate the contributions of FAPAR being cause, effect, or indirect indicator of the vegetations primary production and it is likely that the proportions differ between ecosystem types. Clarifying the relationship between the JRC-FAPAR and GPP using high resolution data has the potential to improve the understanding of landscape scale ecosystem functioning and will be addressed in follow-up studies.

There are several possible reasons why the cumulative FAPAR of the growing season is a poor predictor for annual GPP of deciduous broadleaf forests. The sampled environmental gradient for deciduous broadleaf forests may be too small since the flux sites of these forests are concentrated in the temperate zone with a relatively narrow range of GPP. The absence of a relationship of GPP with absorbed radiation and strong relationship of GPP with IWA implies that GPP of deciduous broadleaf forests is controlled by water stress determined variations of radiation use efficiency. Possibly, rather high frequency variability of radiation use efficiency resulting from

strong stomata activity in response to variations of moisture may control annual GPP, which is not consistently sensed by the FAPAR with 10 day resolution.

	IWA	MAT	RAD	CUM GSL FAPAR	ARAD
ALL	n.s.	0.23	0.07	0.47	0.49
Herbaceous	n.s.	0.61	n.s.	0.74	0.71
Evergreen Forests	n.s.	0.26	0.14	0.69	0.67
Mixed Forest	n.s.	0.58	n.s.	0.53	n.s.
Deciduous Broadleaf Forest	0.55	n.s.	0.16	n.s.	n.s.

Table 5-3: Pearson's correlation (R2) between GPP and the index of water availability (IWA), mean annual temperature (MAT), the annual sum of incoming shortwave radiation (RAD), the cumulative FAPAR of the growing season (cum GSL FAPAR), and absorbed radiation (ARAD). N.s. denotes not significant correlations. IWA is defined as the ratio of actual to potential evapotranspiration. ARAD is calculated as the sum of the product of FAPAR and radiation. IWA, MAT, and RAD data are based on measurements at the tower sites.

5.3 Up-scaling GPP to Europe and corroboration with independent models

5.3.1 Materials and Methods

The up-scaling to the European domain is based on 10 day composite maps of the SeaWiFS FAPAR from 1998-2005 with a spatial resolution of 0.25° in conjunction with the established relationships between the cumulative growing season FAPAR and annual GPP. Firstly, we calculate the cumulative growing season FAPAR on an annual basis for each 0.25° grid cell. Subsequently, we transform the cumulative growing season FAPAR to GPP using the empirical equations. We generate two realisations of European GPP: (1) using the generic function which includes all ecosystem types (FPA), and (2) using separate functions for herbaceous vegetation and evergreen forests in conjunction with a land cover map (FPA+LC). In the latter case we calculate a weighted average GPP, the weights being the land cover fractions within a grid cell:

$$GPP = f_{HERB} \times GPP_{HERB} + f_{EFOREST} \times GPP_{EFOREST} + f_{OTHER} \times GPP_{GENERIC} \qquad \text{Equation (5-6)}$$

where f_{HERB} is the fraction of herbaceous vegetation (grassland + cropland), GPP_{HERB} is GPP as calculated from the equation for herbaceous vegetation, $f_{EFOREST}$ is the fraction of evergreen forests (evergreen coniferous + evergreen broadleaf forest), $GPP_{EFOREST}$ is GPP as calculated

from the equation for evergreen forests, f_{OTHER} is the fraction of other vegetation (here shrub land + deciduous broadleaf forest), $GPP_{GENERIC}$ is GPP as calculated by the generic function that includes all vegetation types. The vegetation fractions were derived from the land cover map of Jung et al., 2006. The three fractions, f_{HERB}, $f_{EFOREST}$, and f_{OTHER} sum up to the total fraction of vegetated land surface for each grid cell. In the following we discuss primarily FPA+LC since accounting for land cover specific relationships should improve the result. We keep FPA to evaluate the impact of additional land cover input.

LPJmL, MOD17+, and ANN simulations

LPJ is a dynamic global vegetation model (DGVM) and originates from the BIOME model family (Haxeltine and Prentice, 1996; Prentice et al., 1992). It simulates the distribution of plant functional types, and cycling of water and carbon on a quasi-daily time-step. LPJ has been used in numerous studies on responses and feedbacks of the biosphere in the Earth System (e.g. Brovkin et al., 2004; Lucht et al., 2002; Schaphoff et al., 2006; Sitch et al., 2005). The version of LPJ used here has been adapted to account for a realistic treatment of croplands using a crop functional type approach (LPJmL, Bondeau et al., 2007).

ANN is a completely data-oriented modelling approach based on Artificial Neural Networks (ANNs) (Papale and Valentini, 2003, Vetter et al., 2007). ANN was trained separately for different vegetation types with flux measurements, meteorological data, and remotely sensed FAPAR from MODIS (collection 4) covering the following vegetation types: deciduous broadleaf forest (11 sites), evergreen needle leaf forests (15 sites), evergreen broadleaf forests and shrub lands (6 sites), grasslands and wetland (18 sites), croplands (12 sites).

MOD17+ is an extended version of the operational MOD17 GPP and NPP product algorithm of Running et al., 2004 to also calculate terrestrial ecosystem respiration. It is a classic RUE model which calculates APAR from the MODIS FAPAR product and net radiation data, and uses temperature and VPD scalars to reduce vegetation type specific maximum RUE. The parameterization to calculate RUE had been optimized for Europe using data from the CarboEurope flux tower measurement network from 2001 and partly 2002 (Reichstein et al., 2004).

LPJmL, MOD17+, and ANN were run on a 0.25° resolution grid with with model input data provided for CarboEurope (Vetter et al., 2007). Meteorological model input is from a regional climate model (REMO, Jacob and Podzun, 1997) that was driven with NCEP reanalysis (Kalnay et al., 1996) at the boundaries of the European model domain (Feser et al., 2001). The simulations were performed for a recent model intercomparison on the 2003 heat wave anomaly (Vetter et al., 2007) in Europe and are available at http://www.bgc-jena.mpg.de/bgc-systems/projects/ce_i/index.shtml. Details on the modelling protocol are available in Vetter et al., 2007.

5.3.2 Results and Discussion

Mean spatial pattern of GPP

LPJmL, MOD17+, ANN, and FPA-LC show relatively low GPP in the boreal and Mediterranean part of Europe but differ on the region of maximum GPP (Figure 5-5). LPJmL concentrates the region of maximum GPP in Western Europe and displays a relatively sharp boundary at ~15° Longitude with much lower GPP in Eastern Europe. ANN simulates a smoother decline of GPP from western to Eastern Europe, while FPA+LC predict an area with a secondary maximum of GPP east of the Baltic Sea. MOD17+ predicts maximum GPP within a belt between 40 and 45° latitude.

2000-2002 mean ***2003 anomaly***	**LPJmL**	**MOD17+**	**ANN**	**FPA**	**FPA+LC**
LPJmL	1	0.47	0.69	0.63	0.59
MOD17+	*0.53*	1	0.75	0.77	0.76
ANN	*0.63*	*0.54*	1	0.85	0.86
FPA	*0.61*	*0.45*	*0.53*	1	0.92
FPA+LC	*0.60*	*0.44*	*0.53*	*0.98*	1

Table 5-4: Matrix of Pearson's correlation coefficients of spatial GPP patterns as predicted by LPJmL, MOD17+, ANN, FPA, and FPA+LC. Above the diagonal: 2000-2002 mean; below diagonal 2003 anomaly relative to the 2000-2002 mean.

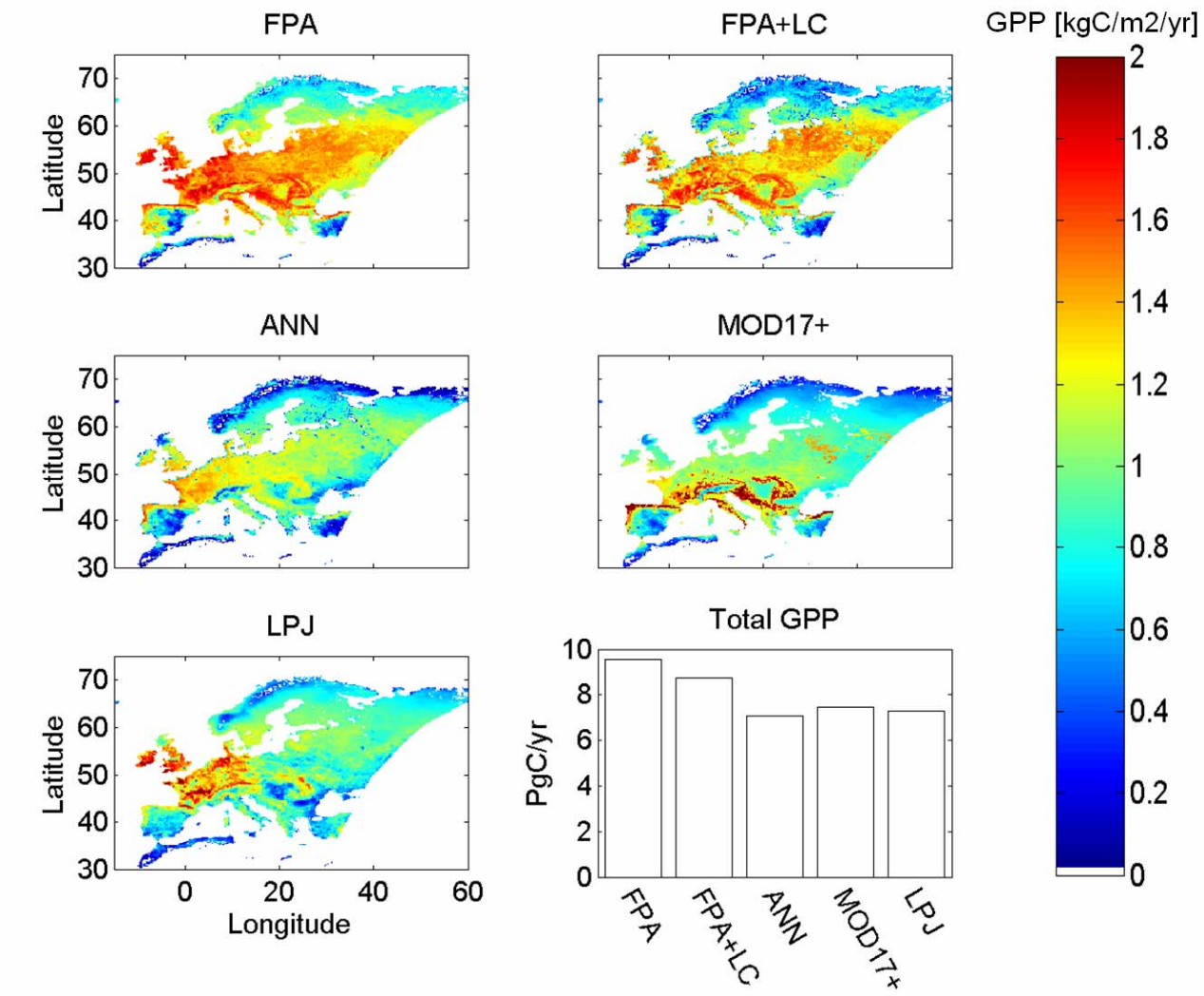


Figure 5-5: Maps of the 2000-2002 mean GPP from LPJmL, ANN, MOD17+, FPA, and FPA+LC. The lower right panel shows the mean total GPP flux over the European domain for the different models.

From a statistical point of view the data oriented models show reasonable good correlation of the spatial pattern among each other, ranging from 0.75 (ANN vs. MOD17+) to 0.86 (ANN vs. FPA+LC, Table 5-4). The spatial correlation of the process model LPJmL with the diagnostic models varies between 0.47 (LPJmL vs. MOD17+) and 0.69 (LPJmL vs. ANN). The strong intercorrelation of mean spatial GPP fields from different data-oriented models indicates an emerging consensus regarding a realistic mean GPP pattern. A close correspondence among these data-oriented model results may be expected given that all are driven by remote sensing input and linked with CarboEuro flux measurements. However, an intercomparison among the different FAPAR data sets (JRC vs. MODIS collection 4), the relationships between the input FAPAR data

and modelled GPP, and correspondence of the different GPP model results shows that MOD17+ and ANN arrive at similar GPP pattern as FPA+LC despite large differences among the spatial FAPAR fields from JRC and MODIS (Figure 5-6). Strong discrepancy of the JRC-FAPAR and an independent approach with FAPAR from MODIS has also been described by Pinty et al., 2007 for several sites. The correlation between MODIS-FAPAR and ANN and MOD17+ GPP patterns is small (R^2 between 0.07 and 0.27) which indicates that the additional meteorological input for ANN, and MOD17+ compensates for the differences of the FAPAR data sets. Therefore, the close agreement among the data-oriented models regarding mean GPP patterns does not result from the fact they are driven by remote sensing input since the weight in MOD17+ and ANN is largely on the meteorology and /or land cover. This underlines the independence between the FPA+LC approach and ANN/MOD17+ and gives some confidence that the consensus regarding the mean GPP pattern is not an artefact of remote sensing input.

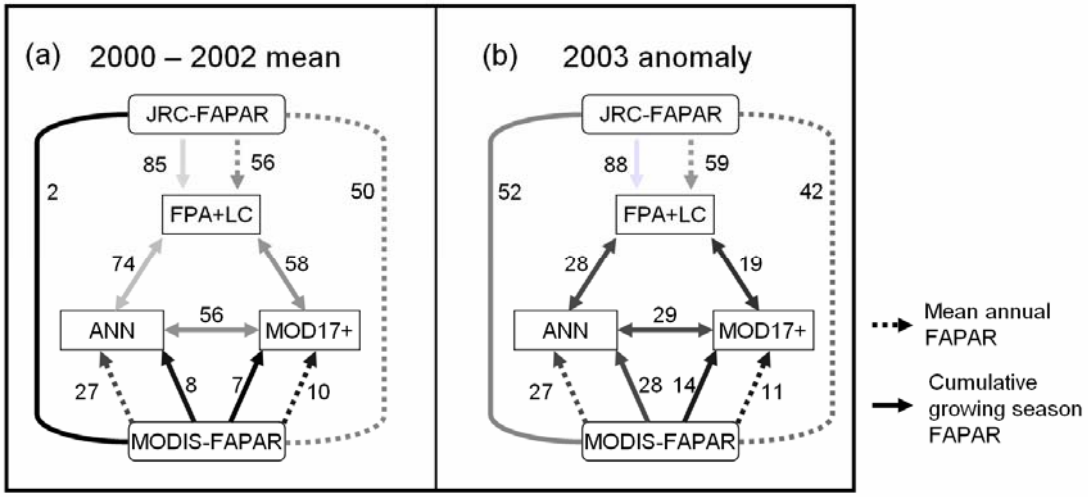


Figure 5-6: Intercomparison of spatial patterns of JRC-FAPAR, MODIS-FAPAR, and GPP estimates from FPA+LC, ANN, and MOD17+ for (a) the 2000-2002 mean, and (b) the 2003 anomaly. The numbers next to the arrows are the variances that are explained by the R^2 (%). The colour of the arrows is scaled in proportion to the R^2. The correlation where FAPAR is one partner is calculated using the mean annual FAPAR (dashed line) and the cumulative growing season FAPAR (solid line). The analysis reveals large differences among the 2000-2002 mean FAPAR fields from JRC and MODIS, while the GPP estimates based on the two different remote sensing products by the different models is less different. Regarding the 2003 anomaly the GPP patterns from the different models are more different than the FAPAR anomalies from JRC and MODIS. It highlights the effect of meteorological input data for MOD17+ and ANN in creating a similar average GPP pattern as FPA+LC despite discrepant remote sensing products, and in creating different anomaly patterns due to model specific sensitivity to meteorological input data.

However, the deviations among MOD17+ and ANN regarding the mean spatial pattern of GPP highlights the uncertainties of modelling GPP over large regions that originates from the modelling approach itself (i.e. 'model structure'). Both used identical driver data including the FAPAR product from MODIS and both being informed by CarboEurope flux tower measurement data. Thus, model-structure seems an equally important source of uncertainty in diagnostic models as in process-oriented models.

The major differences of the mean spatial pattern of GPP among all models are in Central and Eastern Europe. In principle, MOD17+, ANN, and FPA+LC loose some credibility here because no CarboEurope flux stations are present in this region for model tuning; i.e. their predictions have the character of an extrapolation. The region south and east of the Baltic Sea is mainly covered by mixed forests which contains substantial fractions of deciduous broadleaf trees (~30-40%). Since we found no relationship between the cumulative growing season FAPAR and GPP for deciduous forests, we have also limited confidence about the accuracy of FPA and FPA+LC in this region. There are however indications that the pattern of enhanced GPP in this region is realistic. Two other process oriented models, Biome-BGC (Thornton, 1998) and Orchidee (Krinner et al., 2005), place the area of maximum GPP (~1.2 to 1.5 kgC/m2/yr) in Europe south and east of the Baltic sea (Jung et al., 2007c). These two models simulate unreliable results for the parts of Europe that are extensively covered by cropland since they simulate croplands simply as 'productive' grasslands. But for forests, Biome-BGC and Orchidee better reproduce the pattern of changes of forest GPP across Europe than LPJ (Jung et al., 2007a). Simulated GPP for temperate forests from LPJ was found to exhibit a stronger low bias (several hundred grams of carbon per square meter and year) than Biome-BGC and Orchidee. LPJmL also simulates relatively low GPP (800-1000 gC/m2/yr) in the mid-latitude cropland belt in Central and Eastern Europe, which is at least partly related to the fact that irrigation was deactivated in these runs which has an increasing effect with increasing continentality. Moreover, we need to doubt the credibility of the NCEP-REMO meteorological forcing data used to drive LPJmL, MOD17+, and ANN, particularly in Eastern Europe. Regional climate models like REMO have difficulties in predicting the climate accurately in south central and the continental Eastern Europe and tend to produce a dry bias (Hagemann et al., 2004; Jacob et al., 2007). Although the 'true' GPP pattern remains unknown, the above mentioned arguments support the idea that the mean GPP pattern of FPA+LC may display the largest degree of realism among the considered simulations.

Mean total GPP flux of the European domain

In terms of the total flux of GPP over the European domain, LPJmL, MOD17+, ANN, and FPA+LC compare within 1.65 PgC/yr (211 gC/m2/yr). However, MOD17+, ANN and LPJmL simulations are 14 – 19 % lower than FPA+LC (Table 5-5).

	LPJmL	MOD17+	ANN	FPA	FPA+LC
Total GPP 2000-2002 mean [PgC/yr]	7.28	7.47	7.07	9.53	8.72
Mean GPP 2000-2002 mean [gC/m2/yr]	930	954	902	1216	1113
2003 total GPP anomaly [PgC/yr]	-0.64	-0.32	-0.46	-0.33	-0.3
2003 anomaly [%]	-8.84	-4.24	-6.56	-3.44	-3.44

Table 5-5: Total GPP flux of the 2000 and 2002 mean and the 2003 anomaly as predicted by LPJmL, MOD17+, ANN, FPA, and FPA+LC.

A high bias of the cumulative FAPAR based GPP approach is possible if the flux tower sites that were used to calibrate the cumulative FAPAR are biased towards productive rather than average ecosystems. If true, then also MOD17+, and ANN should show a high bias since both have been trained with CarboEurope flux tower measurements too. A low bias of LPJmL may be expected because irrigation was deactivated in this simulation. There is evidence for a likely low biased GPP of MOD17+, ANN and LPJmL simulations induced by the meteorological forcing fields from REMO that were used to drive the models. Jung et al., 2007c have shown that running the Biome-BGC model with an alternative meteorological dataset from ECMWF (ERA 40, ECMWF, 2000) resulted in 20% (1.22 PgC/yr) higher GPP in comparison to NCEP-REMO runs for the same European domain. Zhao et al., 2006 studied the performance of different meteorological driver data from DAO, NCEP, and ECMWF and its effects on global GPP from MOD17 and concluded that ECMWF displays the smallest errors and biases. The impact of biased meteorological input data should be smaller in the case of ANN since the network was trained with the REMO data. However, the result of ANN is affected if the bias changes beyond the region of the distribution of training sites (i.e. especially in Eastern Europe) and is sensitive to what extent the training sites provide a representative and full coverage of climatic conditions. Assuming that the bias resulting from meteorological input as calculated for Biome-BGC for the same European domain (1.22 PgC/yr) is transferable to the other models, all estimates compare within 0.43 PgC/yr or ~5%.

Inter-annual variability

FPA+LC reproduces the well known pattern of the 2003 heat wave anomaly (Ciais et al., 2005; Gobron et al., 2005; Reichstein et al., 2007a; Vetter et al., 2007) in Europe with strong declines of GPP in France and Germany (Figure 5-7). The LPJmL simulations show a very similar pattern but the anomaly extents further east towards the Black Sea. The GPP anomaly from MOD17+ and ANN displays a patchier pattern in western and central Europe and ANN also shows a strong decline near the Black Sea. It is not clear to what extent the strong anomaly near the Black Sea as simulated by LPJmL and ANN is also an artefact of the meteorological input data from NCEP-REMO. By comparing inter-annual variations of GPP in the same European domain as simulated by the Biome-BGC model with REMO and ECMWF meteorology Jung et al., 2007c have shown that south central and eastern Europe was a hot spot of disagreement of GPP inter-annual variability between REMO and ECMWF model runs. Interannual GPP variations simulated due to only different meteorological input were in fact uncorrelated in this region. In general, carbon cycle models are very sensitive to their meteorological forcing fields, which are associated with rather large uncertainties too, particularly related to moisture (precipitation, vapour pressure deficit) and radiation conditions.

FPA+LC also predicts a stronger positive GPP anomaly in northern Europe. The temperature limited forests in northern Europe benefited from the higher temperatures in 2003 which resulted in increased productivity (Vetter et al., 2007). The REMO model tends to produce a warm bias in northern Europe in summer (Hagemann et al., 2004). This general warm bias may be responsible for the less enhanced GPP in northern Europe in LPJmL, MOD17+ and ANN simulations because this decreases the general temperature limitation of the region.

It is interesting to note that the correlations of the spatial patterns of the 2003 anomaly tends to be larger between LPJmL and the data oriented models (0.53 – 0.63) than among the data oriented models (0.44 – 0.54, Table 5-3). The analysis on the relationships between MODIS and JRC-FAPAR anomalies and GPP anomalies of the data-oriented models shows that MODIS and JRC-FAPAR anomalies are more similar than the GPP anomalies (Figure 5-6b). It highlights once more the effect of meteorological input for ANN and MOD17+ and the model specific sensitivity to meteorology in creating GPP variations.

Regarding the total GPP flux anomaly in 2003 relative to the 2000-2002 mean, all estimates range between -0.3 (FPA+LC) and -0.64 PgC/yr (LPJmL). An overemphasized 2003 anomaly by LPJmL is expected because irrigation was deactivated in these simulations. The challenge of studying interannual variability of GPP is highlighted by the fact that the 2003 anomaly, which can be considered an extreme event, is a relatively small signal of only 5.3 +/- 2.4 % (mean +/- standard deviation, Table 5-4) of the mean GPP of the domain.

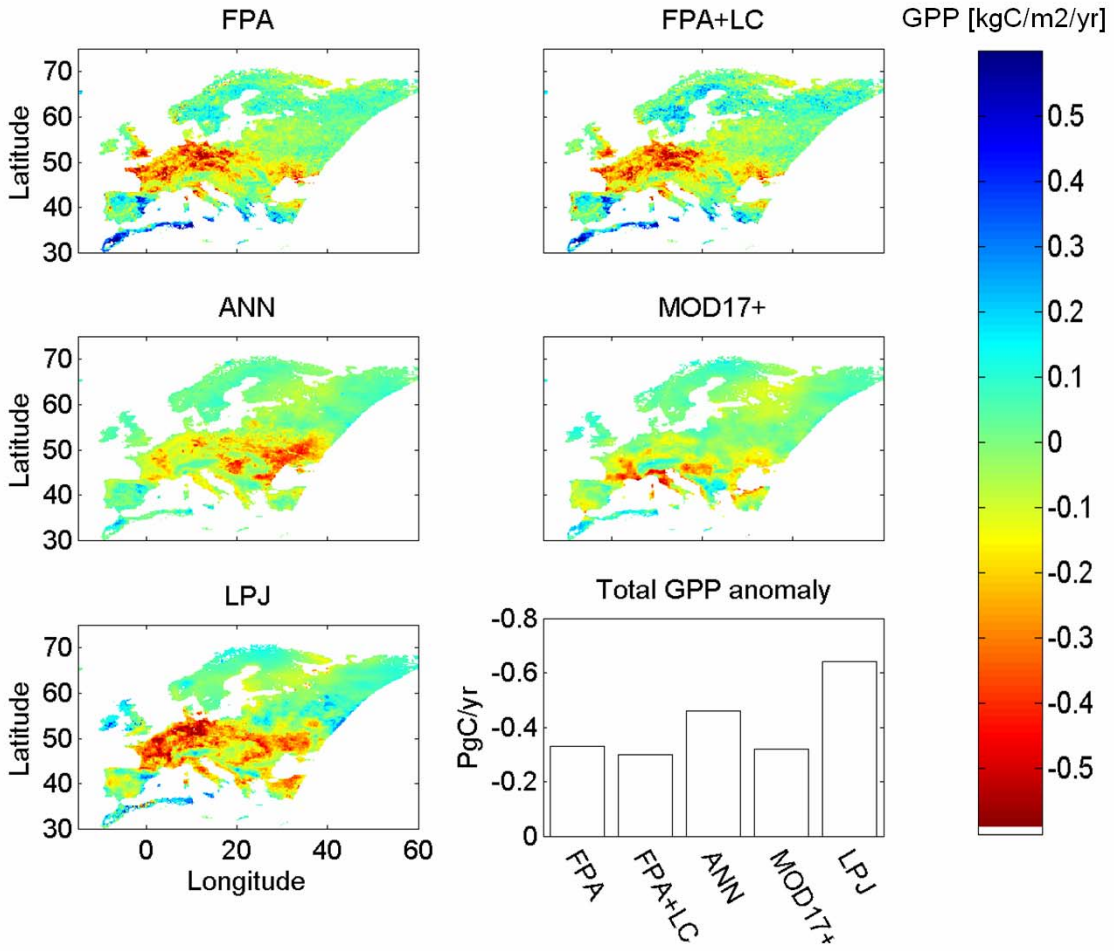


Figure 5-7: Maps of the 2003 anomaly of GPP from LPJmL, ANN, FPA, and FPA+LC. Reference is the 2000-2002 mean. The lower right panel shows the total GPP flux anomaly over the European domain for the different models.

LPJmL, MOD17+, ANN, and FPA+LC agree on most of the main features of inter-annual variability of GPP in four major regions of Europe (Figure 5-8). In Northern Europe which includes the UK and Ireland little variations of GPP are recognizable except a depression in 2001. The positive GPP anomaly in Scandinavia in 2003 is smoothed out due to a negative anomaly in the British Isles. GPP variations in Western Europe are characterised by the marked declines in

2003 and 2005 as predicted by LPJmL, MOD17+, and FPA+LC, while the decrease of GPP from ANN is small in 2003. LPJmL further predicts a very productive 2002 which is not seen in ANN and FPA+LC, and only weakly in MOD17+. The reason why LPJmL may predict a more productive 2002 year may be related to the extensive forest fires in Portugal in 2002 which are only seen by the remotely sensed FAPAR data. A small and a large GPP depression in 2000 and 2003 respectively in central Europe are consistently predicted by the four models. LPJmL and ANN GPP simulations show much larger inter-annual variability in Eastern Europe than FPA+LC. To what extent this phenomenon is related to issues of the meteorological forcing fields of REMO in this region is not clear.

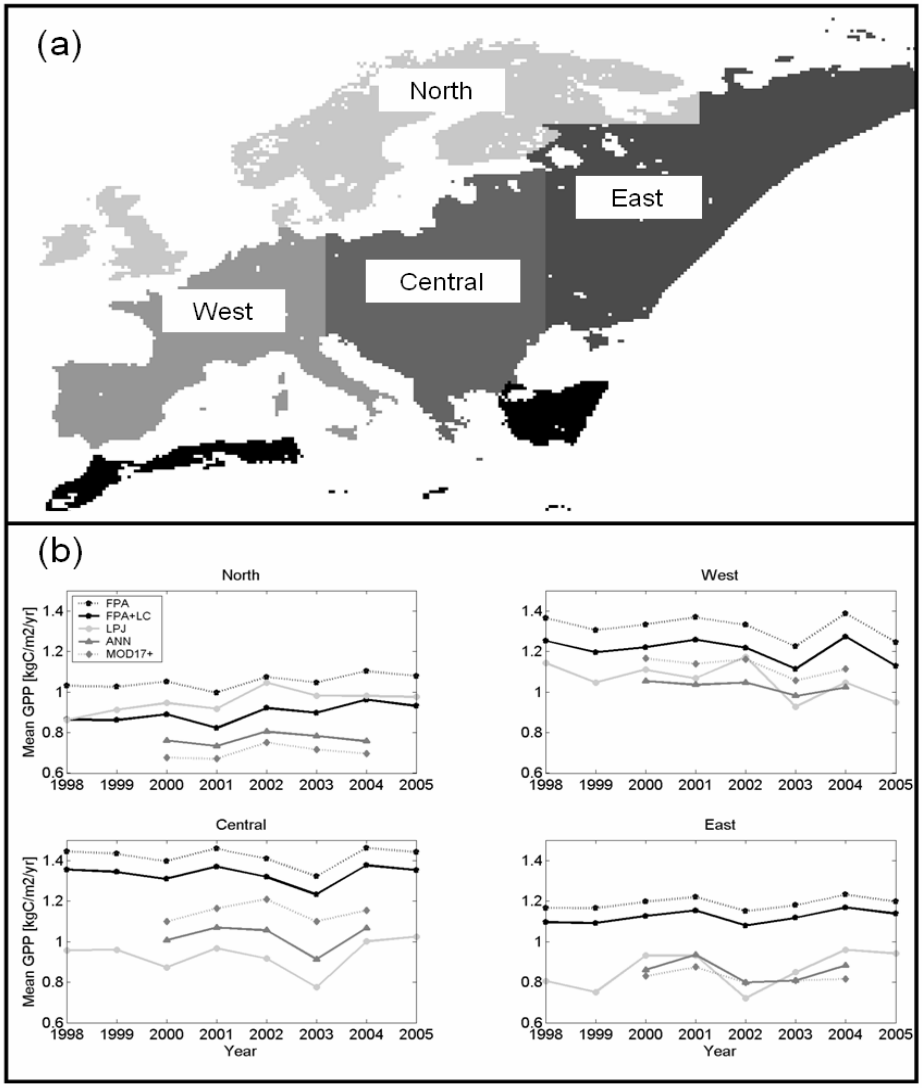


Figure 5-8: (a) defined regions of the European domain, (b) time series of GPP for four major regions as predicted by LPJmL, ANN, FPA, and FPA+LC. Simulations of ANN and MOD17+ are only available from 2000-2004.

On the use of diagnostic models to evaluate prognostic models

The present analysis suggests too little confidence in current diagnostic models to evaluate process-oriented models in terms of interannual GPP variations. To some extent, this may be a detection limit problem given that the very strong anomaly in 2003 is only ~5% of the mean and thus somewhat within the uncertainty. Interannual variability of GPP in Europe is in most parts strongly influenced by variations of plant available water that determine radiation use efficiency. In fact, process models have a theoretical advantage over diagnostic models here as they explicitly simulate soil hydrology and aim to mimic mechanistic plant physiology. In the diagnostic models considered, model specific sensitivity to meteorological input and remote sensing input generates anomaly patterns that are not close enough to each other to justify some confidence in any of the diagnostic models regarding interannual variability. Moreover, interannual variations of simulated GPP are strongly influenced by meteorological input data (MOD17+, ANN), which have been shown to be a large source of uncertainty in particular for interannual GPP variations. To what extent the remotely sensed FAPAR picks-up interannual variations of GPP consistently is not clear. For instance, FAPAR should be very sensitive to water stress effects for herbaceous vegetation, which respond fast by yellowing or senescence. For trees, the FAPAR would sense changes of leaf colour (e.g. yellowing) or leaf shedding and is therefore probably only sensitive to water stress above a certain threshold. However, both herbaceous and tree species are generally present within a pixel (although to varying extents) so that the direction of change should be correct but the magnitude of GPP change less certain.

Emerging consensus among data-oriented models regarding the mean annual spatial pattern of GPP suggests that some diagnostic models may be used to evaluate prognostic models. Caution is necessary given that two of the three considered data-driven models were much closer (ANN vs FPA+LC) than with respect to the third (MOD17+). Hence, a comparison of independent diagnostic models should be conducted to gain some confidence before confronting process-based models with estimates from data-driven models. Spatially varying bias of meteorological input data can further lead to spatially biased GPP estimates.

We conclude that the mean annual GPP pattern from FPA-LC is an accurate and valuable data set for corroboration with results from process models because (1) the independent ANN model arrives at a similar spatial pattern, (2) it identified bias resulting from meteorological input data in

the other models that has been shown in a previous study, and (3) root mean square error of GPP prediction are low and lower than what is reported in the literature for other data-oriented models and process-based models.

5.4 Conclusions

We have shown that the cumulative FAPAR of the growing season derived from space is directly linked to gross carbon uptake in ecosystems in Europe with the exception of deciduous broadleaf forests. The relationship of the two quantities is very strong for herbaceous vegetation and evergreen forests (R^2 of 0.8 and 0.71 respectively) and the associated prediction error for GPP is of the order of 250gC/m2/yr. Given that herbaceous vegetation together with evergreen forests cover ~80 % of the vegetated land surface of Europe we can accurately predict annual GPP of Europe using remotely sensed FAPAR.

By corroborating the FAPAR based GPP against simulations of the LPJmL biosphere model, the radiation use efficiency model MOD17+, and an artificial neuronal network approach on the continental scale we find that the FAPAR based GPP estimates shows reliable spatial and inter-annual variations of GPP. By accounting for bias resulting from the meteorological input data used to drive MOD17+, ANN, and LPJmL the four models compare within ~5% of mean annual GPP of the European domain (8.29 to 8.72 PgC/yr). Our analysis suggests that current data-driven models may be used to evaluate prognostic models regarding mean annual GPP pattern but not regarding interannual variations of GPP, where the uncertainties are possibly larger for diagnostic models.

Uncertainties due to meteorological input data and model structure constitute the largest uncertainties of existing models. A major advantage of the FAPAR based GPP product is that it circumvents both major sources of uncertainty. Given that the FAPAR based GPP approach is entirely based on observed data in conjunction with its accurate performance, it is a valuable tool to quantify GPP over large regions and to evaluate biosphere models. Biosphere models need to be evaluated on their scale of application, i.e. continental to global for which suitable measurement data sets of carbon fluxes are yet lacking. The FAPAR based GPP product closes this gap.

6 Summary, conclusions, and final remarks

6.1 What are the major sources of uncertainties of process-oriented modelling of GPP for Europe?

Uncertainties from model structure are found to exceed uncertainties resulting from input data. However, the accuracy of meteorological forcing fields remains a substantial problem. The effect of different meteorological input data on interannual variations is particularly strong and comparable to the effect of using a different model. In terms of magnitude and spatial pattern of simulated GPP, meteorological input data can introduce substantial bias. Uncertainties resulting from soil input data have not been tested but are expected to be smaller than uncertainties from meteorological input. The effect of different land cover data sets is found to be negligible in comparison uncertainties from meteorology and model structure. Changing the spatial resolution of land cover in the model has a larger effect than changing the land cover data set but this effect is also small.

Essentially three major factors drive the difference among the process-oriented models considered in this study (Biome-BGC, Orchidee, LPJ): (1) the representation of the agricultural sector, (2) the representation of nitrogen dynamics and its interactions with the water and carbon cycle, (3) the representation of carbon-water relations regarding the coupling of canopy conductance and photosynthesis and soil hydrology. A realistic representation of crops as in LPJ is essential to simulate realistic GPP patterns. Representing crops as productive (Orchidee) or fertilized (Biome-BGC) grasslands does not yield comparable/reliable results. The effect of a missing nitrogen cycle in LPJ and Orchidee causes overestimation of leaf area and thus light absorption in nitrogen limited regions like the boreal forest and therefore too large GPP. The model which includes a nitrogen cycle (Biome-BGC) was further found to exhibit different behaviour of interannual variability regarding the relationship and sensitivity to meteorological forcing. The effect of interactions of biogeochemical cycles (here nitrogen-carbon-water) may be an important factor controlling interannual variability, which has received little attention in the past. Modelled water-stress effects on photosynthesis are uncertain and play a particularly large role for interannual variations. Different schemes of coupling between photosynthesis and canopy

conductance, which control transpiration and thus soil hydrology result in different soil water dynamics and sensitivity to soil moisture.

6.2 How realistic are GPP simulations from process-oriented models for Europe?

The comparison of simulated and eddy covariance based GPP and LAI at 37 forest sites across the full climatic gradient running from boreal to Mediterranean climate suggests an overall relative root mean square error of prediction for GPP of 30 % (~420 gC/m2/yr) for Biome-BGC, Orchidee, and LPJ (see chapter 4). Some fraction of this error likely originates from soil and meteorological input data, since the models were not run with in-situ measured meteorology and soil data but with the same gridded input data as the standard CarboEurope simulations to facilitate compatibility. In general, model errors are smallest for temperate forests, and largest for Mediterranean systems, and model errors tend to be smaller for coniferous than for broadleaf forests.

Biome-BGC, Orchidee, and LPJ reproduce qualitatively the broad observed changes of GPP along the gradient of mean annual temperature across Europe: increasing GPP from boreal to temperate climate, and decreasing GPP from temperate to Mediterranean climate. The analysis reveals systematic biases for all models. Orchidee and Biome-BGC tend to slightly better reproduce GPP variations of forests across the continent. Overall, the observed increase of GPP from boreal to temperate environments is too small in the simulations, while the observed decrease of GPP from temperate to Mediterranean climate is too strong in the simulations. The latter cannot be (solely) interpreted in terms of model structural deficiencies due to questionable meteorological driver data but it indicates uncertainties related to modelling water stress effects on photosynthesis.

A method has been developed that allows estimating to what extent inaccurate LAI simulations and thus light absorption cause the too weak gradient of GPP from boreal to temperate forests. This method can be applied consistently for measured and modelled ecosystems, provided that a LAI measurement and a GPP estimate exist for the real forest. The results show that changes of light absorption explain primarily the gradient of GPP from the boreal to the temperate zone.

Since LPJ and Orchidee simulate no change of LAI from boreal to temperate forests they simulate a too weak gradient of GPP. This phenomenon is attributed to missing nitrogen limitation on foliage area in these two models.

The presented analysis is so far the first attempt to evaluate simulated GPP from models that are designed for the large scale actually on their scale of application. However, it falls short in providing information on how realistic the simulations are for the entire continent since crops, which cover ~40 % of the land surface, were not considered due to a lack of data at the time. The comparison of different data-driven GPP simulations with results from LPJ shows reasonable agreement (see Chapter 5). Although LPJ tends to perform not as well as Biome-BGC and Orchidee for forests, LPJ is the only model among the three with an explicit representation of crop functional types, and therefore the only model among the three with reasonable agreement with data-oriented models.

6.3 What is the GPP of Europe?

An empirical GPP model has been constructed using a remotely sensed biophysical vegetation product (JRC-FAPAR) in conjunction with eddy covariance based GPP estimates (see Chapter 5). This new approach allows quantifying annual GPP sum over large regions without the need of additional meteorological data, which are a substantial source of uncertainty. Based on this estimate, which is called FPA+LC, two additional data-driven models (ANN, MOD17+), and LPJ simulations mean annual GPP (2000-2002) can be constrained within ~5% uncertainty. The four model simulations range from 7.07 to 8.72 PgC/yr. By adding the low bias resulting from meteorological input data identified in Chapter 3 for Biome-BGC (1.22 PgC/yr) to the total flux estimates from the model simulation that rely on this meteorological input data set, the models compare within 0.43 PgC/yr (~5%; 8.29 – 8.72 PgC/yr or 1067 – 1113 $gC/m^2/yr$) for the European model domain. This result assumes that the bias from meteorological input on GPP is roughly the same for the different models.

There is further consensus emerging among the data-oriented models regarding the mean spatial pattern of GPP of Europe, particularly among ANN and FPA+LC. Three major arguments

support FPA+LC: (1) the RMSE of the FPA+LC approach is in the order of 250gC/m^2/yr which is substantially lower than what is reported for other data-oriented and process oriented models in the literature (> 388 gC/m^2/yr), (2) the mean spatial pattern of FPA+LC is verified by a neural network upsaling method (ANN, R^2=0.74) using meteorology and MODIS-FAPAR, and (3) the comparison between FPA+LC with ANN, MOD17+, and LPJ uncovers the previously established imprint of biased meteorological input data (see Chaper 3) that were used to drive ANN, MOD17+, and LPJ.

Reliable simulations of interannual variability remain a problem; there is comparatively little consensus among different models, both prognostic and diagnostic, at least quantitatively. To some extent this may be a detection limit problem. For example, the 2003 heat wave anomaly, which can be considered an extreme event, resulted in GPP depression of 3.4 to 8.8 % of the mean of the European domain, depending on the model. Thus, the magnitude of interannual variability is somewhat within the uncertainty of the mean.

6.4 Remarks on evaluations of global terrestrial carbon cycle models

Confronting model simulation with observations allows uncovering uncertainties resulting from model structure and identification of the most adequate model structure among alternative choices. Since data-model comparisons provide information where and how to improve the models, they constitute the first step towards reducing uncertainties and should be regularly repeated. However, sound data-model comparison studies are very demanding and there are no accepted standards of good practise, and also no accepted standards what constitutes a reasonable global terrestrial carbon cycle model. Traditionally, carbon cycle model simulations of NEE at a few FLUXNET sites are compared with eddy covariance measurements of NEE, focusing on how well the model reproduces the seasonal and diurnal cycle of carbon exchange (e.g. Friend et al., 2007; Morales et al., 2005). This model evaluation strategy regarding NEE has several draw backs related to representativity and equifinality. Mean annual NEE and the decadal trend of NEE is largely a function of site history (‘time since disturbance/harvest’) resulting in an imbalance between productivity and respiration (Birdsey et al., 2006; Korner, 2003). Given that relevant information on site history is generally not available the models do not account for this effect. Instead they are brought into long term equilibrium during the spin-up so that the models

can by definition not reproduce mean annual NEE (unless the measured ecosystem is in long-term equilibrium too). Still facilitating data-model comparison by posterior scaling the simulated mean annual NEE to the measured mean annual NEE (e.g. Friend et al., 2007) creates an artificial match between models and data. Testing if the model simulates reasonable differences between day/night and summer/winter is certainly not a rigorous evaluation. Diurnal and seasonal cycles are the major source of variance in the data so that good relationships between observed and modelled data are easily created. Model testing at a limited number of sites allows little judgement of the systematics or generalisation of possible inadequate model structure. The model may seem to work well at some sites but not at other but what is the reason for that? Testing models developed for the large scale only on a few sites with a focus on short-term variations of the diurnal and seasonal cycle is conceptually inadequate. Models developed for continental to global applications should be evaluated on this scale. Comparing simulated and measured NEE is prone to the problem of equifinality. If there is a good match between simulations and measurements of NEE, there is good chance that it appears for the wrong reason. For instance, the CarboEurope model intercomparison on the 2003 heat wave revealed a reasonable correspondence of different models regarding the NEE anomaly, while this was generated by partly different processes in the models (GPP vs. TER, Vetter et al., 2007). Accordingly, if there is consistent mismatch between simulated and observed NEE it is hardly possible to infer where the model is deficient given that NEE is the net effect of a number of processes.

The above mentioned arguments call for a community effort to develop a best practise protocol with certain standards for model evaluations using FLUXNET data. Some suggestions for data-model comparison studies for global biosphere models are: (1) including as many sites as possible to provide a reasonable representative sampling of the environment, (2) investigating systematic data-model mismatches along spatial and temporal environmental gradients (e.g. temperature or moisture availability) that point to model structure deficiencies, (3) constraining the equifinality problem by evaluating additional variables like GPP and LAI and developing methods that allow to identify the importance of e.g. a wrong simulated LAI on simulated GPP, (4) assisting simulations of the full carbon budget from a site specific calibrated stand scale model, (5) forcing the models with in-situ measured meteorological and soil data to avoid confounding effects between input data and model structure issues. If in-situ driver data are not

available (as usually is the case with soil data), sensitivity studies should be performed to estimate an error bar for the simulations that result from uncertain input.

The FLUXNET activity provides a good opportunity to establish a platform that provides easily accessible and pre-processed model driver, flux measurements with uncertainties according to standardized procedures, and ancillary information. In principle, thorough model testing can be made much easier and realised, possibly even with online tools. Model driver data can be downloaded from a website, model simulations uploaded with on the fly output of relevant statistics and visualization. Such strategy has already been successfully employed in benchmarking studies of radiation transfer models (Wildowski et al., 2007).

Confronting GPP simulations of prognostic models with those of diagnostic models are an other option to evaluate prognostic models given that diagnostic models are constrained by 'observed' vegetation properties from remote sensing and tuned with flux estimates from eddy-covariance sites. Diagnostic models bridge the gap between point flux data and grid cell based simulations for large areas from process models. The first comparison among three data oriented models (ANN, MOD17+, FPA+LC) and with one process-based model (LPJ) for GPP presented in chapter 5 reveals optimism but also caution in the use of diagnostic models as benchmark for prognostic models. Optimism results from the fact that (1) the degree of similarity of the mean GPP pattern among the three data oriented models (R^2 between 0.56 and 0.74) is larger than the degree of similarity between the data-oriented model simulations and the process-model's simulation (R^2 between 0.22 and 0.48), and (2) two very independent data-oriented models (ANN, FPA+LC) show a very similar mean GPP pattern ($R^2 = 0.74$). Caution arises because (1) two data-oriented models which are driven by essentially the same input data (ANN, MOD17+) show considerably different mean spatial patterns ($R^2 = 0.56$), indicating that model structure uncertainties are playing an equally important role also for data oriented models, (2) the degree of similarity among all data-oriented models regarding interannual variations is very small (R^2 of 2003 anomaly patterns < 0.3) and smaller than the similarity between the diagnostic and the prognostic model (R^2 of 2003 anomaly patterns < 0.4), suggesting that simulated interannual variability of GPP by data-oriented models is at least equally uncertain as in process-oriented models, (3) two of the three diagnostic models (ANN, MOD17+) require meteorological input data, which introduce substantial uncertainty, and (4) differences of remote sensing based

FAPAR products from MODIS and SEAWiFS (JRC) are very large, which is a key input to diagnostic models. Accepting that interannual variations of GPP are too a large extent controlled by interannual variations of plant available water of the growing season in most parts of Europe, process models that explicitly simulate soil hydrology have a theoretical advantage over diagnostic models, where soil moisture is not or only coarsely represented. In principle, diagnostic models should capture most of the effect of water stress on the vegetation via the remotely sensed FAPAR but it is shown that simulated GPP by ANN and especially MOD17+ is only little influenced by the FAPAR ($R^2 < 0.28$) indicating that input meteorology (radiation, temperature, vapour pressure deficit) drives GPP in these two models.

6.5 Towards reducing uncertainties of global terrestrial carbon cycle models

Uncertainty in carbon cycle modelling can be reduced by improving input data and model structure, and by optimising parameters. Improved meteorological reanalysis have large potential to reduce uncertainties of carbon cycle simulations but this relies on progress primarily outside the carbon community. An example of how input data sets can be improved has been exemplified in chapter 2. An algorithm based on fuzzy logic has been developed that blends different land cover data sets into a new global 1km product with a classification scheme suitable for carbon cycle modelling. The approach exploits similarities among input land cover data sets and thereby minimizes discrepancy between the new SYNMAP data set and input land cover products, while allowing a user defined classification legend for the generated map. However, the generation of SYNMAP as an improved land cover data set for carbon cycle studies does not reduce uncertainties of carbon cycle modelling substantially since non-land cover factors are found to be much more important at least in Europe. SYNMAP has been welcome by the community and is used in recent studies in carbon cycle research (Ahmadov et al., 2007; Churkina et al., 2007; Jung et al., 2007b; Jung et al., 2007c; Pieterse et al., 2007; Vetter et al., 2007, Gerbig et al., 2007). The approach of data-fusion presented in chapter 2 is transferable to any other thematic maps based on categories.

Parameter optimisation has not been investigated in this thesis but receives currently large attention. Such model tuning does only improve the predictability of a model if model structure is

adequate. Moreover, there are a number of key physiological and allometric parameters that should not be represented as a plant functional type specific constant in the models given that changing environmental conditions will lead to different plant characteristics (also within PFTs) that control biogeochemical cycling of the vegetation. Several studies are showing that the variation of plant characteristics can be understood (and modelled) as an optimisation of the vegetation to environmental conditions (e.g. Anten, 2005; Hikosaka, 2005; Shipley et al., 2006). Implementing such principals on optimality in global biosphere models allows moving beyond PFTs with constant characteristics and therefore has potential to improve the predictability in particular for simulations on large time scales (decades, centuries, millennia).

Improving the model structures regarding some deficiencies identified in this thesis is already in progress such as implementing nitrogen dynamics, incorporating crops, and improving soil processes. However, ecosystem modelling is facing the problem of modelling systems from its processes that are largely not understood in a mechanistic sense (e.g. canopy conductance, soil carbon dynamics, interactions of biogeochemical cycles). First principals do not really exist in this field. In contrast to atmospheric or ocean dynamics, the dynamics of ecosystems cannot be described by physical laws and conservation of energy, mass, and momentum. Ecosystem models constitute hypothesis on the functioning of ecosystems. Therefore, a focus must be a much more rigorous testing of the hypothesis expressed by the models than is currently the practise. Such iterative strategy of model development and rigorous evaluation will further enlighten our understanding of ecosystem dynamics.

Global process-oriented terrestrial ecosystem models are in a comparatively early stage. Given the progress of global circulation models during the last 30 years, there is hope for rapidly improving terrestrial ecosystem models in the next decades too. Exploiting diagnostic modelling approaches in the mean time has potential to guide process-oriented modelling. True progress relies on integrated projects of global observational networks, theoreticians, specialists, and modellers. Integration relies on inter-disciplinary communication and understanding. Inter-disciplinary communication and understanding calls for training (some) young scientists across disciplines.

References

Ahlqvist, O., Keukelaar, J. and Oukbir, K., 2003. Rough and fuzzy geographical data integration. International Journal of Geographical Information Science, 17(3): 223-234.

Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neininger, B., Dolman, A. and Sarrat, C., 2007. Mesoscale covariance of transport and CO2 fluxes: evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model. Journal of Geophysical Research-Atmospheres, in press.

Anten, N.P.R., 2002. Evolutionarily stable leaf area production in plant populations. Journal of Theoretical Biology, 217(1): 15-32.

Anten, N.P.R., 2005. Optimal photosynthetic characteristics of individual plants in vegetation stands and implications for species coexistence. Annals of Botany (London), 95(3): 497-508.

Aubinet, M., Berbigier, P., Bernhofer, C., Cescatti, A., Feigenwinter, C., Granier, A., Gruenwald, T., Havrankova, K., Heinesch, B., Longdoz, B., Marcolla, B., Montagnani, L. and Sedlak, P., 2005. Comparing CO2 Storage and Advection Conditions at Night at Diferent Carboeuroflux Sites. Boundary-Layer Meteorology, 116(1): 63-93.

Bartalev, S., Belward, A., Erchov, D.V. and Isaev, A.S., 2003. A new SPOT4-VEGETATION derived land cover map of Eurasia. International Journal of Remote Sensing, 24: 1977-1982.

Bartholome, E. and Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing, 26(9): 1959-1977.

Bartholomé, E. and Belward, A.S., in press. GLC2000: a new approach to global land cover mapping from Earth Observation data. International Journal of Remote Sensing.

Beer, C., 2005. Climatic causes of the evolution of biomass of Russian forests between 1981 and 1999 - An analysis by using a dynamical global vegetation model which is adopted to the boreal zone and constrained by satellite products, Friedrich-Schiller-Unicersitaet, Jena.

Beer, C., Reichstein, M., Ciais, P., Farquhar, G.D. and Papale, D., 2007. Mean annual GPP of Europe derived from its water balance. Geophysical Research Letters, 34(5).

Berry, S.L. and Roderick, M.L., 2002. Estimating mixtures of leaf functional types using continental-scale satellite and climate data. Global Ecology and Biogeography, 11: 23-40.

Birdsey, R., Pregitzer, K. and Lucier, A., 2006. Forest Carbon Management in the United States: 1600-2100. J Environ Qual, 35(4): 1461-1469.

Bondeau, A., Kicklighter, D.W. and Kaduk, J., 1999. Comparing global models of terrestrial net primary productivity (NPP): importance of vegetation structure on seasonal NPP estimates. Global Change Biology, 5: 35-45.

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, Müller, C., Reichstein, M. and Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology, 13(3): 679-706.

Box, E.O., Holben, B.N. and Kalb, V., 1989. Accuracy of the AVHRR vegetation index as a predictor of biomass, primary productivity, and net CO2 flux. Vegetatio, 80: 71-89.

Bradley, B.A., Jacob, R.W., Hermance, J.F. and Mustard, J.F., 2007. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. Remote Sensing of Environment, 106(2): 137-145.
Breda, N., 2003. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. Journal of Experimental Botany, 54(392): 2403-2417.
Brovkin, V., Sitch, S., von Bloh, W., Claussen, M., Bauer, E. and Cramer, W., 2004. Role of land cover changes for atmospheric CO2 increase and climate change during the last 150 years. Global Change Biology, 10(8): 1253-1266.
Cao, M.K., Prince, S.D., Small, J. and Goetz, S.J., 2004. Remotely sensed interannual variations and trends in terestrial net primary productivity 1981-2000. Ecosystems, 7(3): 233-242.
CEOS, in press. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment Of Global Land Cover Maps.
Churkina, G., Brovkin, V., von Bloh, W., Trusilova, K., Jung, M. and Dentener, F., 2007. Synergy of rising nitrogen depositions and atmospheric CO2 on land carbon uptake offsets global warming. Global Change Biology, submitted.
Churkina, G. and Running, S.W., 1998. Contrasting climatic controls on the estimated productivity of global terrestrial biomes. Ecosystems, 1(2): 206-215.
Churkina, G., Tenhunen, J., Thornton, P., Falge, E.M., Elbers, J.A., Erhard, M., Grunwald, T., Kowalski, A.S., Rannik, U. and Sprinz, D., 2003. Analyzing the ecosystem carbon dynamics of four European coniferous forests using a biogeochemistry model. Ecosystems, 6(2): 168-184.
Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A.D., Friedlingstein, P., Grunwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J.M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J.F., Sanz, M.J., Schulze, E.D., Vesala, T. and Valentini, R., 2005. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. Nature, 437(7058): 529-533.
Cihlar, J., 2000. Land cover mapping of large areas from satellites: status and research priorities. International Journal of Remote Sensing, 21(6-7): 1093-1114.
Cleland, E., Chuine, I., Menzel, A., Mooney, H. and Schwartz, M., 2007. Shifting plant phenology in response to global change. Trends in Ecology & Evolution, 22(7): 357-365.
Cook, E.A., Iverson, L.R. and Graham, R.L., 1989. Estimating forest productivity with thematic mapper and biogeographical data. Remote Sensing of Environment, 28: 131-141.
Cowling, S.A. and Field, C.B., 2003. Environmental control of leaf area production: Implications for vegetation and land-surface modeling. Global Biogeochemical Cycles, 17(1).
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., 2001. Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. Global Change Biology, 7(4): 357-373.
Cramer, W., Kicklighter, D.W., Bondeau, A., Moore, B., Churkina, C., Nemry, B., Ruimy, A. and Schloss, A.L., 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. Global Change Biology, 5: 1-15.
Danko, D.M., 1992. The Digital Chart of the World Project. Photogrammetric Engineering and Remote Sensing, 58(8): 1125-1128.

de Rosnay, P. and Polcher, J., 1998. Modelling root water uptake in a complex land surface scheme coupled to a GCM. Hydrology and Earth System Sciences, 2(2-3): 239-255.
DeFries, R.S., Townshend, J.R.G. and Hansen, M.C., 1999. Continuous fields of vegetation characteristics at the global scale at 1-km resolution. Journal of Geophysical Research-Atmospheres, 104(D14): 16911-16923.
Denman, K.L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P.M., Dickinson, R.E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P.L., Wofsy, S.C. and Zhang, X., 2007. Couplings Between Changes in the Climate System and Biogeochemistry. In: S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (Editor), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Desai, A., Richardson, A.D., Moffat, A., Kattge, J., Hollinger, D., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M. and Stauch, V., 2007. Cross site evaluation of eddy covariance GPP and RE decomposition techniques. Agricultural and Forest Meteorology, submitted.
Di Gregorio, A. and Jansen, L.J.M., 2000. Land Cover Classification System (LCCS): Classification Concepts and User Manual. FAO.
Duchemin, B., Goubier, J. and Courrier, G., 1999. Monitoring Phenological Key Stages and Cycle Duration of Temperate Deciduous Forest Ecosystems with NOAA/AVHRR Data. Remote Sensing of Environment, 67(1): 68-82.
Ducoudre, N.I., Laval, K. and Perrier, A., 1993. Sechiba, a New Set of Parameterizations of the Hydrologic Exchanges at the Land Atmosphere Interface within the Lmd Atmospheric General-Circulation Model. Journal of Climate, 6(2): 248-273.
ECMWF, 2000. ERA-40 project plan, Reading, UK.
Eva, H., Belward, A., Miranda, E.E., Bella, C.M., Gond, V., Huber, O., Jones, S., Sgrenzaroli, M. and Fritz, S., 2004. A land cover map of South America. Global Change Biology, 10: 731-744.
Feser, F., Weisse, R. and von Storch, H., 2001. Multi-decadal Atmospheric Modeling for Europe Yields Multi-purpose Data. EOS Transactions, 82: 305-310.
Field, C., 1991. Ecological scaling of carbon gain to stress and resource availability. Response of Plants to Multiple Stresses. Academic Press, San Diego.
Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1): 185-201.
Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F. and Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment, 83(1-2): 287-302.
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., 2006a. Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. Journal of Climate, 19(14): 3337-3353.
Friedlingstein, P., Cox, P.M., Betts, R.A., 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.J., Reick, C.,

Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A.J., Yoshikawa, C. and Zeng, N., 2006b. Climate-Carbon Cycle Feedback Analysis: Results from C4MIP <odel Intercomparison. Journal of Climate, 19: 3337-3353.

Friend, A.D., Arneth, A., Kiang, N.Y., Lomas, M., Ogee, J., Rodenbeckk, C., Running, S.W., Santaren, J.D., Sitch, S., Viovy, N., Woodward, F.I. and Zaehle, S., 2007. FLUXNET and modelling the global carbon cycle. Global Change Biology, 13(3): 610-633.

Fritz, S., Bartholomé, E., Belward, A., Hartley, A., Stibig, H.-J., Eva, H., Mayaux, P., Bartalev, S., Latifovic, R., Kolmert, S., Roy, P.S., Agrawal, S., Bingfang, W., Wenting, X., Ledwith, M., Pekel, J.-P., Giri, C., Mücher, S., Badts, E., Tateishi, R., Champeaux, J.-L. and Defourny, P., 2003. Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version). European Commission, Joint Research Centre (JRC), Ispra.

Fritz, S. and See, L., 2005. Comparison of land cover maps using fuzzy agreement. International Journal of Geographical Information Science, 19(7): 787-807.

GCOS, 2004. Implementation plan for the Global Observing System for Climate in support of the UNFCCC. 1219, WMO, Geneva.

GEOSS, 2005. The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan and Reference Document.

Gerbig, C., Körner, S. and Lin, C.J., 2007. Vertical mixing in atmospheric tracer transport models: error characterization and propagation. Atmos. Chem. Phys. Discuss., 7: 13121-13150.

Giri, C., Zhu, Z. and Reed, B., 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sensing of Environment, 94: 123-132.

Gobron, N., Pinty, B., Aussedat, O., Chen, J., Cohen, W.B., Fensholt, R., Gond, V., Huemmrich, K.F., Lavergne, T., Melin, F., Privette, J.L., Sandholt, I., Taberner, M., Turner, D.P., Verstraete, M. and Wildowski, J.-L., 2006. Evaluation FAPAR products for different canopy radiation transfer regimes: Methodology and results using JRC products derived from SeaWiFS against ground-based estimations. Journal of Geophysical Research-Atmospheres, 111.

Gobron, N., Pinty, B., Aussedat, O., Taberner, M., Faber, O., Mélin, F., Lavergne, T., Robustelli, M. and P., S., 2007. Uncertainty estimates for the FAPAR operational products derived from MERIS - Impact of top-of-atmosphere radiance uncertainties and validation with field data. Remote Sensing of Environment, submitted.

Gobron, N., Pinty, B., Melin, F., Taberner, M., Verstraete, M., Belward, A., Lavergne, T. and Wildowski, J.-L., 2005. The state vegetation in Europe following the 2003 drought. International Journal Remote Sensing Letters, 26(9): 2013-2020.

Gobron, N., Pinty, B., Verstraete, M. and Wildowski, J.-L., 2000. Advanced Vegetation Indices Optimized for Up-Coming Sensors: Design, Performance and Applications. IEEE Transactions on Geoscience and Remote Sensing, 38: 2489-2505.

Goetz, S.J. and Prince, S.D., 1999. Modelling Terrestrial Carbon Exchange and Storage: Evidence and Implications of Functional Convergence in Light-use Efficiency. Advances in Ecological Research, 28: 57-91.

Goetz, S.J., Prince, S.D., Small, J. and Gleason, A.C.R., 2000. Interannual variability of global terrestrial primary production: Results of a model driven with satellite observations. Journal of Geophysical Research-Atmospheres, 105(D15): 20077-20091.

Goward, S.N., Tucker, C. and Dye, D., 1985. North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. Vegetatio, 64: 3-14.

Grace, J., Nichol, C., Disney, M., Lewis, P., Quaife, T. and Bowyer, P., 2007. Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? Global Change Biology, 13: 1484-1497.

Hagemann, S., Machenhauer, B., Jones, R., Christensen, O.B., Deque, M., Jacob, D. and Vidale, P.L., 2004. Evaluation of water and energy budgets in regional climate models applied over Europe. Climate Dynamics, 23: 547-567.

Hansen, M.C., Defries, R.S., Townshend, J.R.G. and Sohlberg, R., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. International Journal of Remote Sensing, 21(6-7): 1331-1364.

Hansen, M.C. and Reed, B., 2000. A comparison of the IGBP DISCover and University of Maryland 1km global land cover products. International Journal of Remote Sensing, 21(6-7): 1365-1373.

Haxeltine, A. and Prentice, I.C., 1996. BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability and competition among plant functional types. Global Biogeochemical Cycles, 10: 693–710.

Haxeltine, A., Prentice, I.C. and Creswell, D.I., 1996. A coupled carbon and water flux model to predict vegetation structure. Journal of Vegetation Science, 7(5): 651-666.

Herold, M., Woodcock, C., Di Gregorio, A., Mayaux, P., A., B., Latham, J. and Schmullius, C.C., in press. A joint initiative for harmonization and validation of land cover datasets. IEEE Transactions on Geoscience and Remote Sensing.

Herold, M., Woodcock, C.E., di Gregorio, A., Mayaux, P., Belward, A.S., Latham, J. and Schmullius, C.C., 2006. A joint initiative for harmonization and validation of land cover datasets. Ieee Transactions on Geoscience and Remote Sensing, 44(7): 1719-1727.

Hicke, J.A., 2005. NCEP and GISS solar radiation data sets available for ecosystem modeling: Description, differences, and impacts on net primary production. Global Biogeochemical Cycles, 19(2).

Hicke, J.A., Asner, G.P., Randerson, J.T., Tucker, C., Los, S., Birdsey, R., Jenkins, J.C., Field, C. and Holland, E., 2002. Satellite-derived increases in net primary productivity across North America, 1982-1998. Geophysical Research Letters, 29(10).

Hikosaka, K., 2005. Leaf Canopy as a Dynamic System: Ecophysiology and Optimality in Leaf Turnover.10.1093/aob/mci050. Ann Bot, 95(3): 521-533.

IGBP-DIS, 2000. Global Soil Data Products CD-ROM. Global Soil Data Task 2000.

Jacob, D., Bärring, L., Christensen, O.B., Christensen, J.H., de Castro, M., Deque, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellström, E., Lenderink, G., Rockel, B., Sanchez, E., Schär, C., Seneviratne, S.I., Somot, S., van Ulden, A. and van den Hurk, B., 2007. An inter-comparison of regional climate models for Europe: model performance in present-day climate. Climatic Change, 81: 31-52.

Jacob, D. and Podzun, R., 1997. Sensitivity studies with the regional climate model REMO. Meteorology and Atmospheric Physics, 63(1-2): 119-129.

JRC, 2003. Global Land Cover 2000 database. European Commission, Joint Research Centre, 2003.

Jung, M., Henkel, K., Herold, M. and Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment, 101(4): 534-553.

Jung, M., Le Maire, G., Zaehle, S., Luyssaert, S., Vetter, M., Churkina, C., Ciais, P., Viovy, N. and Reichstein, M., 2007a. Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe. Biogeosciences Discussion, 4: 1353-1375.

Jung, M., Verstraete, M., Gobron, N., Reichstein, M., Papale, D., Bondeau, A., Robustelli, M. and Pinty, B., 2007b. Diagnostic Assessment of European Gross Primary Production. Global Change Biology, submitted.

Jung, M., Vetter, M., Herold, M., Churkina, G., Reichstein, M., Zaehle, S., Cias, P., Viovy, N., Bondeau, A., Chen, Y., Trusilova, K., Feser, F. and Heimann, M., 2007c. Uncertainties of modelling GPP over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models. Global Biogeochemical Cycles, accepted.

Jung, M., Vetter, M., Herold, M., Churkina, G., Reichstein, M., Zaehle, S., Cias, P., Viovy, N., Bondeau, A., Chen, Y., Trusilova, K., Feser, F. and Heimann, M., in review. Uncertainties of modelling GPP over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models. Global Biogeochemical Cycles.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society, 77(3): 437-471.

Kaplan, J.O., Bigelow, N.H., Prentice, I.C., Harrison, S.P., Bartlein, P.J., Christensen, T.R., Cramer, W., Matveyeva, N.V., McGuire, A.D., Murray, D.F., Razzhivin, V.Y., Smith, B., Walker, D.A., Anderson, P.M., Andreev, A.A., Brubaker, L.B., Edwards, M.E. and Lozhkin, A.V., 2003. Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections. Journal of Geophysical Research-Atmospheres, 108(D19).

Kimball, J.S., Keyser, A.R., Running, S.W. and Saatchi, S.S., 2000. Regional assessment of boreal forest productivity using an ecological process model and remote sensing parameter maps. Tree Physiology, 20(11): 761-775.

Kimball, J.S., Running, S.W. and Saatchi, S.S., 1999. Sensitivity of boreal forest regional water flux and net primary production simulations to sub-grid-scale land cover complexity. Journal of Geophysical Research-Atmospheres, 104(D22): 27789-27801.

Kimball, J.S., White, M.A. and Running, S.W., 1997. BIOME-BGC simulations of stand hydrologic processes for BOREAS. Journal of Geophysical Research-Atmospheres, 102(D24): 29043-29051.

Kirschbaum, M.U.F., Simioni, G., Medlyn, B. and McMurtrie, R.E., 2003. On the importance of including soil nutrient feedback effects for predicting ecosystem carbon exchange. Functional Plant Biology, 30: 223-237.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R. and Fiorino, M., 2001. The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. Bulletin of the American Meteorological Society, 82(2): 247-267.

Knorr, W. and Heimann, M., 1995. Impact of drought stress and other factors on seasonal land biosphere CO2 exchange studied through an atmospheric tracer transport model. Tellus Series B-Chemical and Physical Meteorology, 47.

Knorr, W. and Heimann, M., 2001. Uncertainties in global terrestrial biosphere modeling, part II: Global constraints for a process-based vegetation model. Global Biogeochemical Cycles, 15(1): 227-246.

Knorr, W. and Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biology, 11(8): 1333-1351.

Korner, C., 2003. ATMOSPHERIC SCIENCE: Slow in, Rapid out--Carbon Flux Studies and Kyoto Targets. Science, 300(5623): 1242-1243.

Kramer, K., Leinonen, H., Bartelink, H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, A., Froer, O., Gracia, C.A., Granier, A., Grünwald, T., Hari, P., Jans, W., Kellomarki, S., Loustau, D., Magnani, F., Markkanen, G., Mohren, G.M.J., Moors, E., Nissinen, A., Peltola, H., Sabate, S., Sanchez, A., Sontag, M., Valentini, R. and Vesala, T., 2002. Evaluation of six process-based forest growth models using eddy-covariance measurements of CO2 and H2O fluxes at six forest sites in Europe. Global Change Biology, 8(3): 213-230.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I.C., 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Global Biogeochemical Cycles, 19(1).

Kucharik, C.J., Barford, C.C., El Maayar, M., Wofsy, S.C., Monson, R.K. and Baldocchi, D.D., 2006. A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecological Modelling, 196(1-2): 1-31.

Latifovic, R. and Olthof, I., 2004. Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. Remote Sensing of Environment, 90: 153-165.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21(6-7): 1303-1330.

Lucht, W., Prentice, I.C., Myneni, R.B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W. and Smith, B., 2002. Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. Science, 296(5573): 1687-1689.

Luyssaert, S., Inglima, I., Jung, M., Reichstein, M., Papale, D., Piao, S., Schulze, E.-D., Wingate, L., Matteucci, G., Aubinet, M., Beer, C., Bernhofer, C., Black, K.G., Bonal, D., Chambers, J., Ciais, P., Davis, K.J., Delucia, E.H., Dolman, A., Don, A., Gielen, B., Grace, J., Granier, A., Grelle, A., Griffis, T., Grünwald, T., Guidolotti, G., Hanson, P., Harding, R., Hollinger, D., Kolari, P., Kruijt, B., Kutsch, W., Lagergren, F., Laurila, T., Law, B., Le Maire, G., Lindroth, A., Magnani, F., Marek, M., Mateus, J., Migliavacca, M., Misson, L., Montagnani, L., Moncrieff, J., Moors, E., Munger, J.W., Nikinmaa, E., Loustau, D., Pita, G., Rebmann, C., Richardson, A.D., Roupsard, O., Saigusa, N., Sanz, M., Seufert, G., Soerensen, L., Tang, J., Valentini, R., Vesala, T. and Janssens, I.A., accepted. The CO2-balance of boreal, temperate and tropical forests derived from a global database. Global Change Biology.

Luyssaert, S., Janssens, I.A., Sulkava, M., Papale, D., Dolman, A.J., Reichstein, M., Hollmen, J., Martin, J.G., Suni, T., Vesala, T., Loustau, D., Law, B.E. and Moors, E.J., 2007. Photosynthesis drives anomalies in net carbon-exchange of pine forests at different latitudes. Global Change Biology, 0(ja).

Magnani, F., Mencuccini, M., Borghetti, M., Berbigier, P., Berninger, F., Delzon, S., Grelle, A., Hari, P., Jarvis, P.G., Kolari, P., Kowalski, A.S., Lankreijer, H., Law, B.E., Lindroth, A., Loustau, D., Manca, G., Moncrieff, J.B., Rayment, M., Tedeschi, V., Valentini, R. and Grace, J., 2007. The human footprint in the carbon cycle of temperate and boreal forests. Nature, 447(7146): 849-851.

Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G. and Minerbi, S., 2005. Importance of advection in the atmospheric CO2 exchanges of an alpine forest. Agricultural and Forest Meteorology, 130(3-4): 193-206.

Mayaux, P., Bartholome, E., Fritz, S. and Belward, A., 2004. A new land cover map of Africa for the year 2000. Journal of Biogeography, 32: 861-877.

Mayaux, P., Eva, H., Gallego, J., Strahler, A.H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E.E., Di Bella, C.M., Ordoyne, C., Kopin, Y. and Roy, P.S., 2006. Validation of the global land cover 2000 map. Ieee Transactions on Geoscience and Remote Sensing, 44(7): 1728-1739.

Mayaux, P., Strahler, A., Eva, H., Herold, M., Shefali, A., Naumov, S., Dorado, A., Di Bella, C., Johansson, D., Ordoyne, C., Kopin, I., Boschetti, L. and Belward, A., in press. Validation of the Global Land Cover 2000 Map. IEEE Transactions on Geoscience and Remote Sensing.

McGuire, A.D., Sitch, S., Clein, J.S., Dargaville, R., Esser, G., Foley, J., Heimann, M., Joos, F., Kaplan, J., Kicklighter, D.W., Meier, R.A., Melillo, J.M., Moore, B., Prentice, I.C., Ramankutty, N., Reichenau, T., Schloss, A., Tian, H., Williams, L.J. and Wittenberg, U., 2001. Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO2, climate and land use effects with four process-based ecosystem models. Global Biogeochemical Cycles, 15(1): 183-206.

Moffat, A., Papale, D., Reichstein, M., Hollinger, D., Richardson, A.D., Barr, A., Beckstein, C., Braswell, B.H., Churkina, G., Desai, A., Falge, E., Gove, J.H., Heimann, M., Hui, D., Jarvis, A., Kattge, J., Noormets, A. and Stauch, V., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agricultural and Forest Meteorology, accepted.

Monsi, M. and Saeki, T., 1953. Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion. Japanese Journal of Botany, 14: 22-52.

Monteith, J.L., 1965. Light distribution and photosynthesis in field crops. Ann Bot, 29: 17-37.

Moorcroft, P.R., 2006. How close are we to a predictive science of the biosphere? Trends in Ecology & Evolution, 21(7): 400-407.

Morales, P., Sykes, M.T., Prentice, I.C., Smith, P., Smith, B., Bugmann, H., Zierl, B., Friedlingstein, P., Viovy, N., Sabate, S., Sanchez, A., Pla, E., Gracia, C.A., Sitch, S., Arneth, A. and Ogee, J., 2005. Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. Global Change Biology, 11(12): 2211-2233.

Muchoney, D., Strahler, A., Hodges, J. and LoCastro, J., 1999. The IGBP DISCover confidence sites and the system for terrestrial ecosystem parameterization: Tools for validating global land-cover data. Photogrammetric Engineering and Remote Sensing, 65(9): 1061-1067.

Nakaji, T., Ide, R., Oguma, H., Saigusa, N. and Fujinuma, Y., 2007. Utility of spectral vegetation index for estimation of gross CO2 flux under varied sky conditions. Remote Sensing of Environment, 109: 274-284.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences, 3(4): 571-583.

Papale, D. and Valentini, A., 2003. A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biology, 9(4): 525-535.

Pieterse, G., Bleeker, A., Vermeulen, A.T., Wu, Y. and Erisman, J.W., 2007. High resolution modelling of atmosphere-canopy exchange of acidifying and eutrophying components and carbon dioxide for European forests. Tellus Series B-Chemical and Physical Meteorology, 59(3): 412-424.

Pinty, B., Lavergne, T., Vossbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M.M. and Widlowski, J.L., 2007. Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products. Journal of Geophysical Research-Atmospheres, 112(D10).

Pinty, B., Leprieur, C. and Verstraete, M., 1993. Towards a Quantitative Interpretation of Vegetation Indices Part 1: Biophysical Canopy Properties and Classical Indices. Remote Sensing Reviews, 7: 127-150.

Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A. and Klooster, S.A., 1993. Terrestrial Ecosystem Production - a Process Model-Based on Global Satellite and Surface Data. Global Biogeochemical Cycles, 7(4): 811-841.

Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A. and Solomon, A.M., 1992. A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate. Journal of Biogeography, 19(2): 117-134.

Prentice, I.C., Heimann, M. and Sitch, S., 2000. The carbon balance of the terrestrial biosphere: Ecosystem models and atmospheric observations. Ecological Applications, 10(6): 1553-1573.

Prince, S.D. and Goward, S.N., 1995. Global primary production: A remote sensing approach. Journal of Biogeography, 22(4-5): 815-835.

Raupach, M.R., Rayner, P.J., Barrett, D.J., DeFries, R.S., Heimann, M., Ojima, D.S., Quegan, S. and Schmullius, C.C., 2005. Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Global Change Biology, 11(3): 378-397.

Rayner, P.J., Scholze, M., Knorr, W., Kaminski, T., Giering, R. and Widmann, H., 2005. Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). Global Biogeochemical Cycles, 19(2).

Reich, P.B. and Oleksyn, J., 2004. Global patterns of plant leaf N and P in relation to temperature and latitude. Proceedings of the National Academy of Sciences of the United States of America, 101(30): 11001-11006.

Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogee, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., Grunwald, 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., 2007a. Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis. Global Change Biology, 13(3): 634-651.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biology, 11(9): 1424-1439.

Reichstein, M., Papale, D., Valentini, R., Aubinet, M., Bernhofer, C., Knohl, A., Laurila, T., Lindroth, A., Moors, E., Pilegaard, K. and Seufert, G., 2007b. Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites. Geophysical Research Letters, 34.

Reichstein, M., Valentini, A., Running, S. and Tenhunen, J., 2004. Improving remote-sensing based GPP estimates (MODIS-MOD17) through inverse parameter estimation with CARBBOEUROPE eddy covariance flux data, EGU. Geophysical Research Abstracts, Nice, pp. 547-560.
Roxburgh, S.H., Barrett, D.J., Berry, S.L., Carter, J.O., Davies, I.D., Gifford, R.M., Kirschbaum, M.U.E., McBeth, B.P., Noble, I.R., Parton, W.G., Raupach, M.R. and Roderick, M.L., 2004. A critical overview of model estimates of net primary productivity for the Australian continent. Functional Plant Biology, 31(11): 1043-1059.
Ruimy, A., Dedieu, G. and Saugier, B., 1996. TURC: A diagnostic model of continental gross primary productivity and net primary productivity. Global Biogeochemical Cycles, 10: 269-285.
Ruimy, A., Kergoat, L. and Bondeau, A., 1999. Comparing global models of terrestrial net primary productivity (NPP): analysis of differences in light absorption and light-use efficiency. Global Change Biology, 5: 56-64.
Running, S., Loveland, T.R., Pierce, L.L., Nemani, R. and Hunt, E.R., 1995. A remote-sensing based vegetation classification logic for global land-cover analysis. Remote Sensing of Environment, 51(1): 39-48.
Running, S.W., 1994. Testing Forest-Bgc Ecosystem Process Simulations across a Climatic Gradient in Oregon. Ecological Applications, 4(2): 238-247.
Running, S.W. and Gower, S.T., 1991. Forest-Bgc, a General-Model of Forest Ecosystem Processes for Regional Applications .2. Dynamic Carbon Allocation and Nitrogen Budgets. Tree Physiology, 9(1-2): 147-160.
Running, S.W. and Hunt, E.R.J., 1993. Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for the global scale. In: J.R. Ehleringer and C.B. Field (Editors), Scaling physiological processes: leaf to globe. Academic Press, San Diego, pp. 141-158.
Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M.S., Reeves, M. and Hashimoto, H., 2004. A continuous satellite-derived measure of global terrestrial primary production. Bioscience, 54(6): 547-560.
Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N. and Ohno, H., 2005. A crop phenology detection method using time-series MODIS data. Remote Sensing of Environment, 96(3-4): 366-374.
Scepan, J., 1999. Thematic validation of high-resolution global land-cover data sets. Photogrammetric Engineering and Remote Sensing, 65(9): 1051-1060.
Schaphoff, S., Lucht, W., Gerten, D., Sitch, S., Cramer, W. and Prentice, I.C., 2006. Terrestrial biosphere carbon storage under alternative climate projections. Climatic Change, 74(1-3): 97-122.
Scurlock, J.M.O., Cramer, W., Olson, R.J., Parton, W.J. and Prince, S.D., 1999. Terrestrial NPP: Toward a consistent data set for global model evaluation. Ecological Applications, 9(3): 913-919.
See, L. and Fritz, S., in press. Towards a global hybrid land cover map for the year 2000. International Journal of Geographical Information Science.
Shipley, B., Lechoweicz, M.J., Wright, I. and Reich, P.B., 2006. Fundamental trade-offs generating the worldwide leaf economics spectrum. Ecology, 87(3): 535-541.
Sims, D.A., Rahman, A.F., Cordova, V.D., El-Masri, B.Z., Baldocchi, D.D., Flanagan, L.B., Goldstein, A.H., Hollinger, D.Y., Misson, L., Monson, R.K., Oechel, W.C., Schmid, H.P., Wofsy, S.C. and Xu, L.K., 2006. On the use of MODIS EVI to assess gross primary

productivity of North American ecosystems. Journal of Geophysical Research-Biogeosciences, 111(G4).
Sitch, S., Brovkin, V., von Bloh, W., van Vuuren, D., Assessment, B. and Ganopolski, A., 2005. Impacts of future land cover changes on atmospheric CO2 and climate. Global Biogeochemical Cycles, 19(2).
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., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology, 9(2): 161-185.
Smith, J.H., Wickham, J.D., Stehman, S.V. and Yang, L., 2002. Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy. Photogrammetric Engineering and Remote Sensing, 68: 65-70.
Strahler, A.H., Muchoney, D., Borak, J., Friedl, M., Gopal, S., Lambin, E. and Moody, A., 1999. MODIS Land Cover Product: Algorithm Theoretical Basis Document (ATBD) Version 5.0, Boston.
Tedeschi, L.O., 2006. Assessment of the adequacy of mathematical models. Agricultural Systems, 89: 225-247.
Thornton, P., 1998. Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets. PhD Thesis, University of Montana, Missoula.
Thornton, P., 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology, 113: 185-222.
Trusilova, K., Churkina, G., Vetter, M., Reichstein, M., Schumacher, J., Knohl, A., Rannik, Ü., Grünwald, T., Moors, E. and Granier, A., in review. Parameter estimation for the terrestrial ecosystem model BIOME-BGC using nonlinear inversion. Ecosystem Modelling.
Turner, D.P., Cohen, W.B. and Kennedy, R.E., 2000. Alternative spatial resolutions and estimation of carbon flux over a managed forest landscape in Western Oregon. Landscape Ecology, 15(5): 441-452.
van Dijk, A., Dolman, A. and Schulze, E.D., 2005. Radiation, temperature, and leaf area explain ecosystem carbon fluxes in boreal and temperate European forests. Global Biogeochemical Cycles, 19(GB2029).
Verstraete, M. and Pinty, B., 1996. Designing optimal spectral inices for remote sensing applications. IEEE Transactions on Geoscience and Remote Sensing, 34: 1254-1265.
Verstraete, M.M., Gobron, N., Aussedat, O., Robustelli, M., Pinty, B., Widlowski, J.-L. and Taberner, M., 2007. An automatic procedure to identify key vegetation phenology events using the JRC-FAPAR products. Advances in Space Research, In Press, Corrected Proof: 273.
Vetter, M., Churkina, G., Bondeau, A., Chen, Y., Ciais, P., Feser, F., Freibauer, A., Geyer, R., Heimann, M., Jones, C., Jung, M., Papale, D., Reichstein, M., Tenhunen, J., Tomelleri, E., Viovy, N. and Zaehle, S., in preparation. A multi-model comparison of growing season carbon flux anomalies:spatial (2003) and temporal variations (1980-2005).
Vetter, M., Churkina, G., Jung, M., Reichstein, M., Zaehle, S., Bondeau, A., Chen, Y., Ciais, P., Feser, F., Freibauer, A., Geyer, R., Jones, C., Papale, D., Heimann, M., Tenhunen, J., Tomelleri, E., Trusilova, K., Viovy, N. and Heimann, M., 2007. Analyzing the causes and spatial pattern of the European 2003 carbon flux anomaly in Europe using seven models. Biogeosciences Discussion, 4: 1201-1240.

Vetter, M., Wirth, C., Bottcher, H., Churkina, G., Schulze, E.D., Wutzler, T. and Weber, G., 2005. Partitioning direct and indirect human-induced effects on carbon sequestration of managed coniferous forests using model simulations and forest inventories. Global Change Biology, 11(5): 810-827.

White, A., Thornton, P., Running, S., Ramakrishna and Nemani, R., 2000. Parameterization and sensitivity analysis of the Biome-BGC terrestrial ecosystem model: Net primary production controls. Earth Interactions, 4(3): 1-85.

White, M.A. and Nemani, R.R., 2006. Real-time monitoring and short-term forecasting of land surface phenology. Remote Sensing of Environment, 104(1): 43-49.

Wildowski, J.-L., Robustelli, M., Disney, M., Gastellu-Etchegorry, J.P., Lavergne, T., Lewis, P., North, P.R.J., Pinty, B., Thompson, R. and Verstraete, M., 2007. The RAMI On-line Model Checker (ROMC): A web-based benchmarking facility for canopy reflectance models. Remote Sensing of Environment, in press.

Wright, I.J., Groom, P.K., Lamont, B.B., Poot, P., Prior, L.D., Reich, P.B., Schulze, E.D., Veneklaas, E.J. and Westoby, M., 2004a. Leaf trait relationships in Australian plant species. Functional Plant Biology, 31(5): 551-558.

Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Warton, D.I. and Westoby, M., 2005. Modulation of leaf economic traits and trait relationships by climate. Global Ecology and Biogeography, 14(5): 411-421.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. and Villar, R., 2004b. The worldwide leaf economics spectrum. Nature, 428(6985): 821-827.

Wright, J.P., Naeem, S., Hector, A., Lehman, C., Reich, P.B., Schmid, B. and Tilman, D., 2006. Conventional functional classification schemes underestimate the relationship with ecosystem functioning. Ecology Letters, 9(2): 111-120.

Xiao, X.M., Hollinger, D., Aber, J., Goltz, M., Davidson, E.A., Zhang, Q.Y. and Moore, B., 2004. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. Remote Sensing of Environment, 89(4): 519-534.

Yang, L., Ichii, K., White, M.A., Hashimoto, H., Michaelis, A., Votava, P., Zhu, A., Huete, A., Running, S. and Nemani, R., 2007. Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through Support Vector Machine Approach. Remote Sensing of Environment, 110: 109-122.

Zaehle, S., Sitch, S., Smith, B. and Hatterman, F., 2005. Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. Global Biogeochemical Cycles, 19(3).

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C. and Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment, 84(3): 471-475.

Zhao, M., Running, S.W. and Nemani, R.R., 2006. Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. Journal of Geophysical Research-Biogeosciences, 111(G1).