Diagnostic assessment of European gross primary production

MARTIN JUNG*†‡, MICHEL VERSTRAETE‡, NADINE GOBRON‡, MARKUS REICHSTEIN*, DARIO PAPALE§, ALBERTE BONDEAU¶, MONICA ROBUSTELLI‡ and BERNARD PINTY‡

*Max Planck Institute for Biogeochemistry, Jena, Germany, †International Max Planck Research School on Earth System Modelling, Hamburg, Germany, ‡European Commission – DG Joint Research Centre, Institute for Environment and Sustainability, Global Environment Monitoring Unit, Ispra (VA), Italy, §Department of Forest Environment and Resources, DISAFRI, University of Tuscia, Viterbo, Italy, ¶Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany

Abstract

We present an approach to estimate gross primary production (GPP) using a remotely sensed biophysical vegetation product (fraction of absorbed photosynthetically active radiation, 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 $250 \, \text{gC m}^{-2} \, \text{yr}^{-1}$. The cumulative growing season FAPAR integrates over a number of effects relevant for GPP such as the length of the growing season, the vegetation's response to environmental conditions and the amount of light harvested that is available for photosynthesis. We corroborate the proposed GPP estimate (noted FAPAR-based productivity assessment + land cover, 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 Lund–Potsdam–Jena managed Land biosphere model (LPJmL). The closest agreement of the mean spatial GPP pattern among the four models is between FPA + LC and ANN ($R^2 = 0.74$). At least some of the discrepancy between FPA-LC and the other models result from biases of meteorological forcing fields for MOD17 +, ANN and LPJmL. Our analysis further implies that meteorological information is to a large degree redundant for GPP estimation when using the JRC-FAPAR. A major advantage of the FPA + LC approach presented in this paper lies in its simplicity and that it requires no additional meteorological input driver data that commonly introduce substantial uncertainty. We find that results from different data-oriented models may be robust enough to evaluate process-oriented models regarding the mean spatial pattern of GPP, while there is too little consensus among the diagnostic models for such purpose regarding inter-annual variability.

Keywords: carbon cycle, eddy covariance, Europe, FAPAR, GPP, model

Received 1 October 2007; revised version received 28 April 2008 and accepted 12 February 2008

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 signifi-

cance of GPP in driving the net carbon balance, both in terms of spatial, as well as temporal variations (Ciais *et al.*, 2005; van Dijk *et al.*, 2005; Luyssaert *et al.*, 2007; Reichstein *et al.*, 2007b). GPP is thus a critical flux that drives the carbon budget of ecosystems.

GPP estimates at the ecosystem level are available from eddy covariance measurements of net ecosystem exchange (NEE) by separating NEE into the gross fluxes GPP and ecosystem respiration (Reichstein *et al.*, 2005; Moffat *et al.*, 2007). GPP assessments on a large scale can

be achieved with a data-oriented approach. Such diagnostic models are driven by spatial fields of meteorological data and remotely sensed vegetation properties. Radiation-use efficiency (RUE) models (Monsi & 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 RUE. RUE is usually calculated as a land cover-specific property that is reduced by scalars according to meteorological or soil hydrological conditions. Up-scaling the carbon fluxes from FLUXNET sites to the continent using artificial intelligence has further been proposed (Papale & 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 by up-scaling the ecosystem's water-use efficiency.

Large-scale patterns of GPP estimated from diagnostic models are in principle also very useful for evaluations of process-oriented models. The major purpose of process-oriented biogeochemical ecosystem models is to run in prognostic mode within an Earth System Model framework. 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.*, 2006). Complementary to testing the process representation of the models on the site level (e.g. Morales *et al.*, 2005), prognostic models developed for the continental to global scale should also be tested as to what extent they are capable of reproducing the major patterns of variability on a large scale, which are provided by the diagnostic models. However, diagnostic models are also subject to uncertainty resulting from the respective approach itself and input data, and the robustness of patterns from different diagnostic models have not yet been evaluated.

One major source of uncertainty for both prognostic and diagnostic models is meteorological forcing field. The choice of the meteorological input dataset alone can result in a 20% difference of simulated GPP as estimated by Jung *et al.* (2007b) for Europe using the Biome-BGC model (Thornton, 1998) and the globe by Zhao *et al.* (2006) with the MOD17 RUE model (Running *et al.*, 2004). More importantly, the choice of meteorological input data can strongly influence the final GPP patterns, in particular regarding inter-annual variations (Jung *et al.*, 2007b).

There is renewed interest in directly relating remotely sensed vegetation properties to GPP (Rahman *et al.*, 2005; Sims *et al.*, 2006b), which circumvents the problem of the meteorological data, and making assumptions on ecosystem functioning in some diagnostic models (e.g. RUE models). Relationships between the integrated

Normalized Difference Vegetation Index (NDVI) over the growing season and net primary productivity (NPP) had already been reported in the 1980s for the regions in North America (Goward *et al.*, 1985; Box *et al.*, 1989; Cook *et al.*, 1989). Recently, Sims *et al.* (2006b) 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 the growing season data points are compared. Advanced remote sensing-based vegetation products, used in conjunction with networks of eddy covariance flux measurement sites and standardized data processing chains (Papale *et al.*, 2006), offer now unprecedented possibilities to investigate and exploit relationships between remotely sensed vegetation properties and gross carbon uptake of ecosystems at the continental scale.

In this study, we develop a simple empirical model to estimate annual sums of GPP over Europe based only on remotely sensed fraction of absorbed photosynthetic active radiation (FAPAR or fPAR) and eddy covariance flux tower measurements. Subsequently, we apply this model to the European domain and corroborate our results with independent simulations from the Lund–Potsdam–Jena managed Land (LPJmL) biosphere model, the RUE model MOD17 + and a neural network-based up-scaling of GPP. This comparison aims to (1) evaluate our proposed approach against state of the art models on the continental level and (2) to infer reasons for similar or dissimilar GPP patterns to gain a better understanding of uncertainties of diagnostic GPP modelling. The latter objective has important implications as to what extent results from current data-oriented GPP models may serve as a reference for process-based models.

Materials and methods

Developing a simple empirical GPP model by linking remotely sensed FAPAR to gross carbon uptake from eddy covariance flux towers

GPP estimates from eddy covariance flux tower measurements. GPP is estimated by separating the measured net flux of carbon from the land surface to the atmosphere (net ecosystem exchange, NEE) into its gross constituent fluxes GPP and terrestrial ecosystem respiration (TER).

$$\text{NEE} = \text{TER} - \text{GPP}. \qquad (1)$$

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.
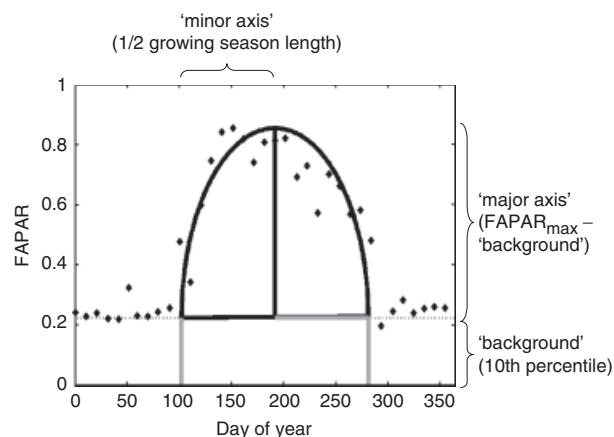
Annual sums of GPP based on flux-separated eddy covariance measurements of NEE are subject to various uncertainties that may be introduced by a number of processing steps: quality control [u*-filtering, spike removal, storage correction (Papale et al., 2006)], gap filling (Moffat et al., 2007) and partitioning of NEE into GPP and TER (Reichstein et al., 2005; Desai et al., 2008). Effects of problematic micrometeorological conditions that are not filtered out by the quality controls 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). Uncertainties are surely site-specific but are usually within $100 \, \mathrm{g \, C \, m^{-2} \, yr^{-1}}$. We follow Reichstein et al. (2007b) who used an uncertainty of $200 \, \mathrm{g \, C \, m^{-2} \, yr^{-1}}$ for annual GPP sums as a conservative estimate.

JRC-FAPAR from the SeaWiFS sensor. FAPAR is a fraction of absorbed radiation in the PAR domain by green vegetation. The JRC (Joint Research Centre) algorithm capitalizes on the physics of remote sensing measurements and minimizes contaminating effects of sun-target-sensor geometry, atmospheric aerosol and soil brightness changes (Gobron et al., 2000). 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 the atmospheric effects on these measurements. In practice, the generic FAPAR algorithm implements a two-step procedure where the spectral bi-directional reflectance functions (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 MEdium Resolution Imaging Spectrometer (MERIS) instrument, show that the impact of the top-of-atmosphere radiance uncertainties on the products is <10% (Gobron et al., 2008). 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 (available from http://fapar.jrc.ec.europa.eu/).

Calculating the cumulative growing season FAPAR (CGS-FAPAR). The information from the FAPAR record that is sensitive to vegetation productivity is the integrated FAPAR of the growing season. Several methods of different complexity and computational demand have been proposed to identify start and end events of the growing season from multi-temporal satellite data (e.g. Duchemin et al., 1999; Zhang et al., 2003; Sakamoto et al., 2005; White & Nemani, 2006; Bradley et al., 2007; Verstraete 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 because 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 the robust performance of a growing season length algorithm.

We developed a simple method that calculates the cumulative growing season FAPAR (CGS-FAPAR) value without determining the start and end events of a growing season explicitly, which is computationally efficient and robust against noise. Firstly, we determine the 'background' that is typical for nongrowing season conditions (Fig. 1). 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 the sites but is rather consistent among the years at one site. Secondly, we subtract the background value from the FAPAR record and sum all the positive values of a year. We analysed the relationship of the sum of FAPAR values above this

Fig. 1 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.

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. The length of the growing season can be approximated by assuming that the FAPAR record is shaped like half of an ellipse [Eqns (2) and (3)] (i.e. similar to bell shaped), which is a reasonable approximation in most cases. By employing this geometrical solution, we do not need to detect start and end events of the growing season. Given the area of the ellipse (twice the accumulated FAPAR above the background) 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 values (represented by the rectangle in Fig. 1) that had been initially subtracted [Eqn (4)]. 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.

$$2 \times \text{CUM}_{\text{BG}} = \pi \times \text{MAX}_{\text{BG}} \times \frac{\text{GSL}}{2} \qquad (2)$$

$$\text{GSL} = \frac{4 \times \text{CUM}_{\text{BG}}}{\pi \times \text{MAX}_{\text{BG}}} \qquad (3)$$
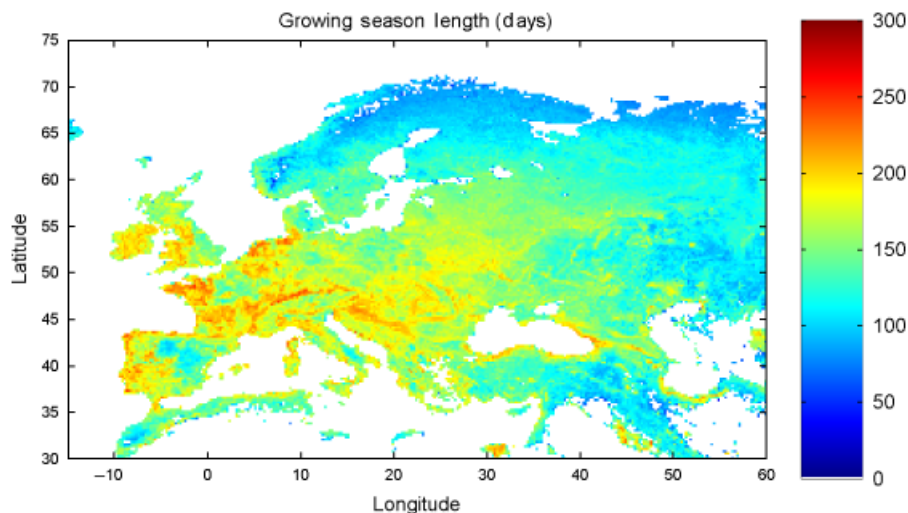
$$\text{CGS} - \text{FAPAR} = \text{CUM}_{\text{BG}} + \text{GSL} \times \text{BG}, \qquad (4)$$

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

Our method of approximating the length of the growing season 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, the uncertainty on the final CGS-FAPAR value is small because 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 values. The retrieved European pattern of growing season length shows the main expected gradients of decreasing growing season length towards the boreal regions, with increasing continentality and with increasing aridity (Fig. 2).

Linking the CGS-FAPAR to GPP at flux tower sites. We regress the calculated CGS-FAPAR at the flux tower sites against annual sums of GPP using all the available site-years and stratified by vegetation types. We use the 10-day composite FAPAR time series that is available for all the CarboEurope sites for the exact pixel and for a 3×3 window. We make use of the latter because this provides a less noisy and a more complete record. We perform gap filling where short gaps of maximum three 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.

Fig. 2 Map of the mean growing season length (1998–2002) based on the proposed algorithm to calculate the cumulative FAPAR of the growing season.
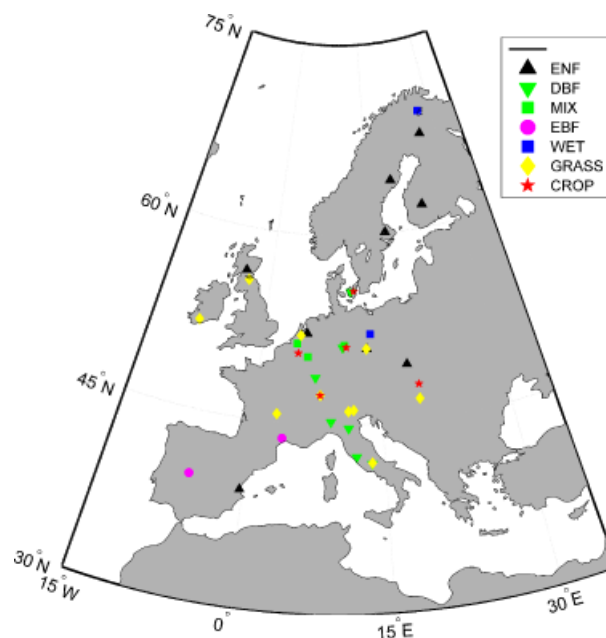
We employ measures of similarity between the (not gap filled) 1×1 and 3×3 extracts FAPAR time series at the sites to assess the degree of local landscape heterogeneity. Large heterogeneity makes it likely that the smaller footprint of the flux tower is not representative for the area sampled by the satellite pixels (i.e. the scale mismatch between the tower and satellite footprint matters here), and such sites should not be included. Sites are excluded where Pearson's correlation coefficient and the modelling efficiency measure (Tedeschi, 2006) of the 1×1 and 3×3 FAPAR record are below 0.7 and 0.5, respectively. We further discard data 2005 onwards because the original sensor radiances are no longer available at the full spatial resolution. This processing yields a set of GPP and CGS-FAPAR data for 39 sites (117 site-years) (Fig. 3). These 39 sites span various vegetation types, as well as a large environmental gradient, from boreal to Mediterranean climates.

Up-scaling GPP to Europe and corroboration with independent models

FAPAR-based productivity assessment (FPA). The up-scaling to the European domain is based on the 10-day composite maps of the SeaWiFS FAPAR from 1998 to 2005 with a spatial resolution of $0.25°$ in conjunction with the established relationships between the CGS-FAPAR and annual GPP (see 'The relationship between the cumulative growing season FAPAR and GPP at flux tower sites'). Firstly, we calculate the CGS-FAPAR on an annual basis for each $0.25°$ grid cell. Subsequently, we transform the CGS-FAPAR to GPP using the empirical equations. We generate two realizations 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) because we find improved relationships for these two vegetation types relative to the generic relationship over all vegetation types (see 'The relationship between the cumulative growing season FAPAR and GPP at flux tower sites'). For FPA + LC, we calculate a weighted average GPP, the weights being the land cover fractions within a grid cell:

$$\text{GPP} = f_{\text{HERB}} \times \text{GPP}_{\text{HERB}} + f_{\text{EFOREST}} \times \text{GPP}_{\text{EFOREST}} + f_{\text{OTHER}} \times \text{GPP}_{\text{GENERIC}}, \tag{5}$$

where f_{HERB} is the fraction of herbaceous vegetation (grassland + cropland), GPP_{HERB} is the GPP as calculated from the equation for herbaceous vegetation, f_{EFOREST} is the fraction of evergreen forests (evergreen coniferous + evergreen broadleaf forest), $\text{GPP}_{\text{EFOREST}}$ is the GPP as calculated from the equation for evergreen forests, f_{OTHER} is the fraction of other vegetation (here shrub land + deciduous broadleaf forest) and $\text{GPP}_{\text{GENERIC}}$

Fig. 3 Map of Europe with CarboEurope sites used in this study. ENF, evergreen needle leaf forest; DBF, deciduous broadleaf forest; MIX, mixed forest; EBF, evergreen broadleaf forest; WET, wetlands; GRASS, grasslands; CROP, croplands. The number of sites and site-years used for each vegetation type are given in Table 1.

is the GPP as calculated by the generic function that includes all vegetation types. The vegetation fractions were derived from the (static) 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. We discuss primarily FPA + LC because accounting for land cover-specific relationships should improve the result, but 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 (Prentice *et al.*, 1992; Haxeltine & Prentice, 1996). It simulates the distribution of plant functional types and cycling of water and carbon on a quasidaily time step. LPJ has been used in numerous studies on responses and feedbacks of the biosphere in the Earth system (e.g. Lucht *et al.*, 2002; Brovkin *et al.*, 2004; Sitch *et al.*, 2005; Schaphoff *et al.*, 2006). 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 & Valentini, 2003; Vetter *et al.*, 2008). 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 (six sites), grasslands and wetland (18 sites) and 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 TER. It is a classic RUE model which calculates APAR from the MODIS FAPAR product and net radiation data and uses temperature, VPD and precipitation-related 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 from 2002 (Reichstein, 2006).

LPJmL, MOD17 + and ANN were run on a 0.25° resolution grid with model input data provided for CarboEurope (Vetter *et al.*, 2008). Meteorological model input is from a regional climate model (REMO; Jacob & 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 inter-comparison on the 2003 heat wave anomaly (Vetter *et al.*, 2008) 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.* (2008).

Results and discussion

The relationship between the cumulative growing season FAPAR and GPP at flux tower sites

The CGS-FAPAR explains more than 50% of the variance in annual GPP data ($R^2 = 0.56$, $n = 117$) across different vegetation types and years (Fig. 4a) when we fit a logarithmic function [Eqn (6)]. While testing the different functions, the logarithmic fit seemed most appropriate. A linear regression would have been possible as well but it decreases model performance in terms of R^2 and root mean square error (RMSE); higher order polynomials provide an apparent higher fit than the logarithmic function but are likely to over-fit the data.

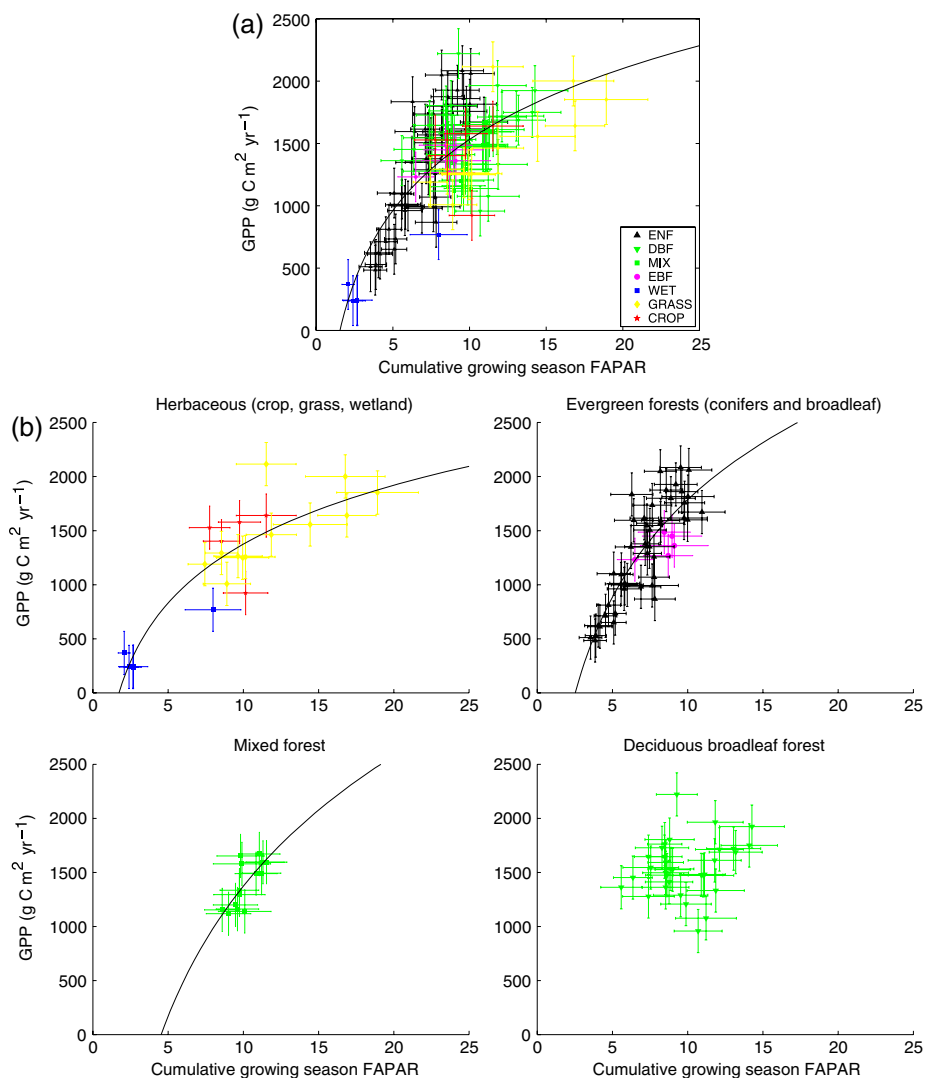$$GPP = a \times \ln(CGS - FAPAR) + b. \qquad (6)$$

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 (Fig. 4b). The relationship becomes substantially stronger for herbaceous ecosystems ($R^2 = 0.8$) and evergreen forests ($R^2 = 0.71$). For mixed forests, the CGS-FAPAR still explains more than

50% of the variation of annual GPP, while we find no significant relationship for deciduous forests (Table 1).

The RMSE is $< 250 \, g \, C \, m^{-2} \, yr^{-1}$ 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 $414–453 \, g \, C \, m^{-2} \, yr^{-1}$ ($n = 37$; 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 $490 \, g \, C \, m^{-2} \, yr^{-1}$ (R^2 between 0.56 and 0.74; Yang *et al.*, 2007) and $388–414 \, g \, C \, m^{-2} \, yr^{-1}$ (R^2 between 0.33 and 0.47) for the support vector machine approach of Yang *et al.* (2007) (Table 2). Reviewing the literature, we noticed that statistics of predictability are commonly reported for the 8-daily values (temporal resolution of MODIS products) but not for the 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 2). The development of our regression model explicitly for the annual time scale is probably an important reason why our RMSEs are comparatively smaller.

We conclude that the relationship between the CGS-FAPAR 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 nonsignificant relationship for deciduous forests is relatively small as the prediction error is still small (due to the small range of GPP data) and also because the 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 1 km land cover map (Jung *et al.*, 2006)].

How are the cumulative FAPAR of the growing season and annual gross carbon uptake linked? There are several reasons why the CGS-FAPAR acts as a good indicator for GPP. The relationship between FAPAR and GPP emerges because canopy foliage is both determinant and consequence of the vegetation's primary production. Firstly, FAPAR determines carbon assimilation because it constrains the amount of light that is available for carboxylation. On top of light absorption, greenness-related remote sensing-based vegetation indexes and the JRC-FAPAR contain a signal of varying chlorophyll content. This traces variations of RUE, for example, seasonally as leafs age and become darker or yellow. The vegetation index EVI (enhanced vegetation index) was found to co-vary with

Fig. 4 Scatter plots of the cumulative growing season FAPAR and GPP for (a) all data points and (b) stratified by ecosystem types. The curves correspond to the best fit. See Table 1 for additional information.

Table 1 Statistics on the relationship between the CGS-FAPAR and annual sums of GPP for different groups of ecosystems

	A	B	R^2	RMSE $(gC\,m^{-2}\,yr^{-1})$	Relative RMSE	Number of sites year	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 forests	–	–	ns	248	0.16	32	8

A and B are the parameters of the logarithmic fit from Eqn (5). All given correlations are highly significant ($P < 0.01$, Pearson's correlation).

ns, not significant ($P > 0.05$).

The relative RMSE is defined as the RMSE divided by mean GPP.

RUE, and this co-variation was found important for the close correspondence of EVI and GPP (Sims *et al.*, 2006b; Nakaji *et al.*, 2007).

Secondly, FAPAR is a consequence of (past) productivity, as it reflects the amount of photosynthetic active tissue of the land surface. In this sense, FAPAR

Table 2 Compilation of RMSE and R^2 values for data-oriented GPP models from multi-site studies using eddy covariance GPP estimates

Model name	Model input	Vegetation types	RMSE 8-daily (gC m⁻² day⁻¹)	R^2 8-daily	RMSE annual (gC m⁻² yr⁻¹)	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	–	–	–	D. Papale (unpublished results)
		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 optimized with Carbo-Europe flux tower measurements	FAPAR (MODIS), radiation (in situ), temperature (in situ), VPD (in situ)	Various	1.49	0.77	–	–	–	M. Reichstein (unpublished results)
		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)
		Nonforest	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	
		Nonforest	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. (2006a, b)
		Cold climate evergreen	–	0.66*	–	–	–	
		West coast evergreen	–	0.20*	–	–	–	
		Deciduos forest	–	0.59*	–	–	–	
		Grassland	–	0.33*	–	–	–	

The R^2 values reported by Sims et al. refer to 16-day averages rather than the 8-day averages and are labelled with an asterisk. The RMSE and R^2 on an annual scale from Yang et al. (2007) were calculated from Table 4 in Yang et al. (2007).

would most closely be related to foliage NPP that in turn is strongly correlated with GPP (Litton et al., 2007). 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. In connection with this point, we need to emphasize the possible role of the herbaceous fraction in the satellite footprint as 'bio-indicators' for the temporal GPP variability. The herbaceous fraction at evergreen forest sites (under storey, surrounding crop/grassland) may enhance the growing season signal (cf. Sims et al., 2006a, b) and act as an indicator for the productivity of the forest. The FAPAR signal of herbaceous vegetation is more sensitive to the variability of environmental conditions because herbaceous plants respond fast to, for example, water stress by yellowing or senescence. Trees in contrast may experience similar stress, which results in reduced photosynthesis due to the closure of stomata but do not necessarily react with leaf yellowing or shedding that the FAPAR would pick up. Also Reichstein et al. (2007a) attributed FAPAR decreases in evergreen needle-leaf forests during the 2003 summer heat wave largely to leaf yellowing of herbaceous plants (understorey, mixed pixels).

There are several possible reasons why the CGS-FAPAR is a poor predictor for annual GPP of the cold-deciduous broadleaf forests. The sampled environmental gradient for deciduous broadleaf forests may be too small because the flux sites of these forests are concentrated in the temperate zone with a relatively narrow range of GPP. Indeed, the RMSE is not much bigger than that for most of the other forest types. The absence of a relationship of GPP with absorbed radiation and a strong relationship of GPP with water availability (ratio of actual to potential evapotranspiration; data not shown) implies that GPP of deciduous broadleaf forests is controlled by water stress-determined variations of RUE. Possibly, rather high-frequency variability of RUE resulting from strong stomata activity in response to variations of moisture may control annual GPP, which is not consistently sensed by the FAPAR. In addition, the signal from the herbaceous fraction may be small because the deciduous broadleaf sites tend to be closed-canopy (beech) systems, or, the signal from the herbaceous fraction may be decoupled from the GPP variability of the forests due to different rooting depths and, thus access to soil moisture. All these possible factors that may eliminate the relationship between the CGS-FAPAR and GPP for deciduous broadleaf sites could also be responsible for some of the scatter seen in the relationship for other vegetation types. We refer the interested reader to Sims et al. (2006a) and Sims et al. 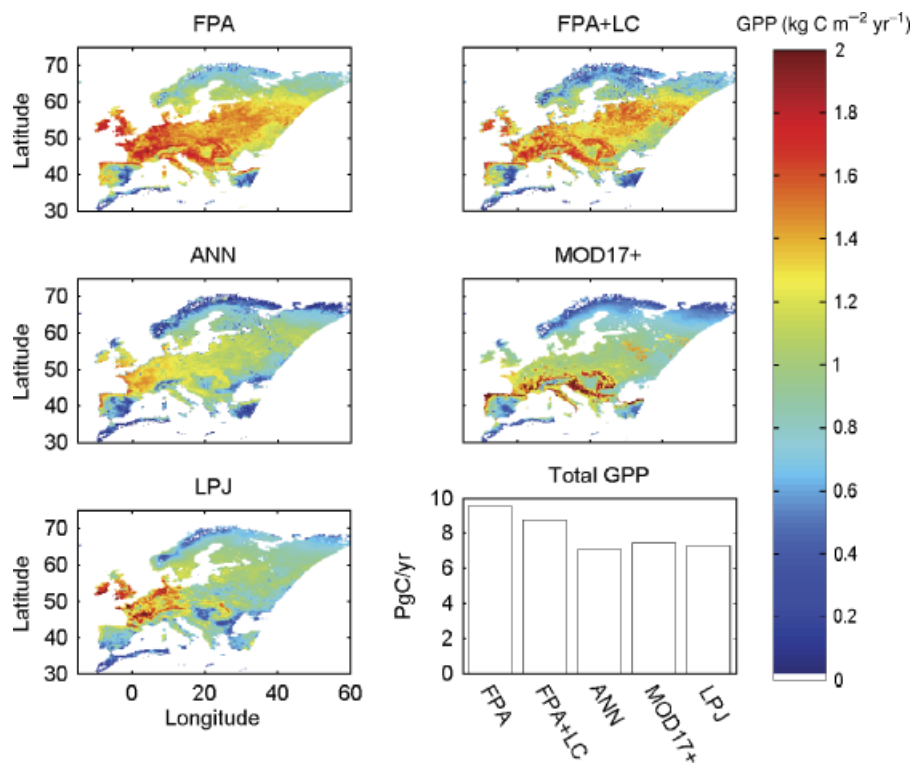(2006b) for further discussions on the relationship between greenness-related remote sensing vegetation products and GPP.

In summary, the CGS-FAPAR can be (1) cause, (2) effect or (3) indirect indicator of GPP. Clearly, we cannot separate the contributions of FAPAR as being the cause, the effect or the indirect indicator of the vegetations primary production and it is likely that the proportions differ between the ecosystem types. Clarifying this using high-resolution data in conjunction with spectrometer measurements can improve diagnostic GPP models and has the potential to improve the understanding of landscape scale ecosystem functioning.

Comparison of GPP patterns from different models on the continental scale

In this section, we present results from the up-scaling to the European domain using the established relationships between the cumulative growing season FAPAR and GPP. We corroborate the results with simulations from two other data-oriented models, ANN and MOD17 +, and the process-oriented LPJmL model. We evaluate the degree of similarity among the different models regarding the mean spatial pattern of GPP, the spatial GPP pattern of the 2003 heat wave anomaly and the magnitude of the total GPP flux over the European domain. This assessment also aims at inferring sources of uncertainty of large-scale GPP modelling.

The estimates of mean total flux of GPP (years 2000–2002) over the European domain range from 7.1 Pg C yr^{-1} or 900 g C m^{-2} yr^{-1} (ANN) to 8.7 Pg C yr^{-1} or 1110 g C m^{-2} yr^{-1} (FPA + LC). LPJmL, MOD17 +, ANN and FPA + LC show a relatively low mean annual GPP in the boreal and Mediterranean part of Europe but differ in the region of maximum GPP (Fig. 5). LPJmL concentrates the region of maximum GPP in Western Europe and displays a relatively sharp boundary at $\sim 15°$ longitude with much lower GPP in Eastern Europe. ANN simulates a smoother decline of GPP from Western to Eastern Europe, while FPA + LC predicts 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. Despite these distinct features, it is worthwhile noting that the mean GPP patterns of the four models presented here are much closer to each other in comparison to results from three process-oriented models shown in Jung et al. (2007b). From a statistical point of view, the data-oriented models show reasonable good correlation of the spatial pattern of mean GPP among each other, ranging from 0.75 (ANN vs. MOD17 +) to 0.86 (ANN vs. FPA + LC; Table 3). The spatial correlation of the process model LPJmL with the diagnostic models is lower; it varies between 0.47 (LPJmL vs. MOD17 +) and 0.69 (LPJmL vs. ANN).

Fig. 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.

Table 3 Matrix of Pearson's correlation coefficients of spatial GPP patterns as predicted by LPJmL, MOD17 + , ANN, FPA and FPA + LC
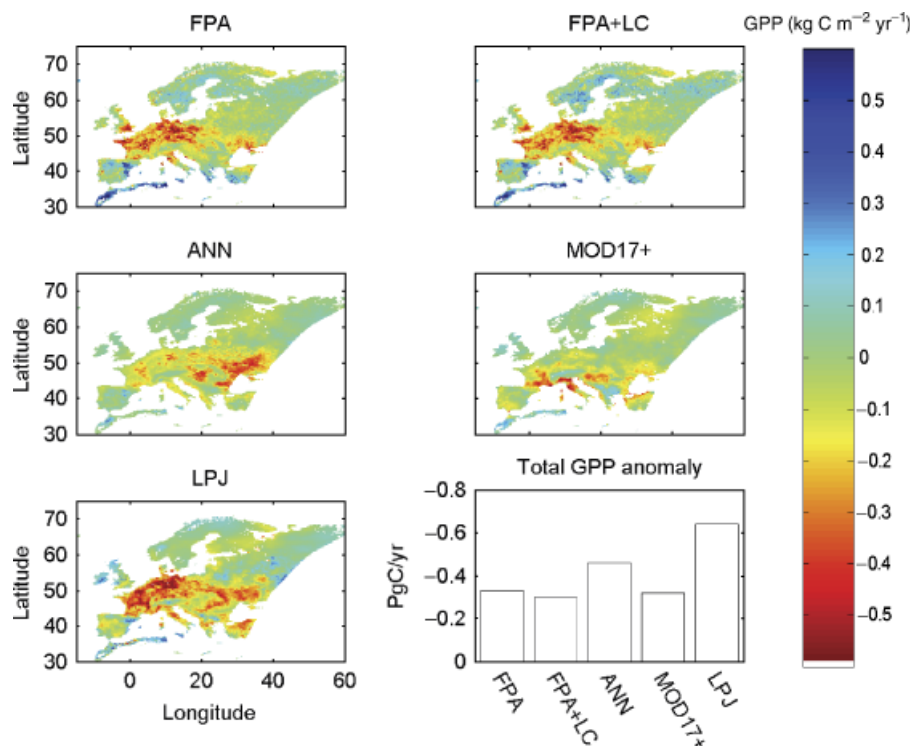
	2000–2002 mean				
2003 anomaly	LPJmL	MOD17 +	ANN	FPA	FPA + LC
LPJmL	1	0.47	0.69	0.63	0.59
MOD17 +	*0.53*	1	0.75	0.77	0.76
ANN	*0.63*	*0.54*	1	0.85	0.86
FPA	*0.61*	*0.45*	*0.53*	1	0.92
FPA + LC	*0.60*	*0.44*	*0.53*	*0.98*	1

Above the diagonal: 2000–2002 mean; below the diagonal 2003 anomaly (numbers set in italics) relative to the 2000–2002 mean.

The imprint of the 2003 heat wave and drought anomaly on the productivity of European ecosystems provides a valuable test case for carbon models in terms of inter-annual variability. Integrated over the European domain, the estimated GPP reduction of 2003 relative to the mean of 2000–2002 ranges from 0.3 (FPA + LC) to 0.64 Pg C (LPJmL). 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.*, 2008) in Europe with strong declines of GPP

in France and Germany (Fig. 6). 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. In general, the correlations of the spatial patterns of the 2003 anomaly among the models are lower (0.44–0.63) than that for the mean GPP pattern (0.47–0.86; Table 3). Interestingly, the correlations of the spatial patterns of the 2003 anomaly tend to be larger between the process model LPJmL and the data-oriented models (0.53–0.63) than among the data-oriented models (0.44–0.54).

Impacts of uncertain meteorological forcing on LPJmL, MOD17 + and ANN. The NCEP-REMO meteorological forcing data used to drive LPJmL, MOD17 + and ANN are associated with substantial uncertainties (Zhao *et al.*, 2006; Chen *et al.*, 2007), and the difference to FPA + LC can to some extent be attributed to issues with the meteorological data. Jung *et al.* (2007b) have shown that running the Biome-BGC model with an alternative meteorological dataset from ECMWF (ERA 40; Uppala *et al.*, 2005) resulted in 20% (1.22 Pg C yr^{-1}) higher GPP in comparison with NCEP-REMO runs for

Fig. 6 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.

the same European domain. Assuming that the bias resulting from meteorological input as calculated for Biome-BGC for the same European domain ($1.22 \, \text{Pg C yr}^{-1}$) is transferable to the other models, the estimates from all the four models compare within $0.43 \, \text{Pg C yr}^{-1}$ or $\sim 5\%$.
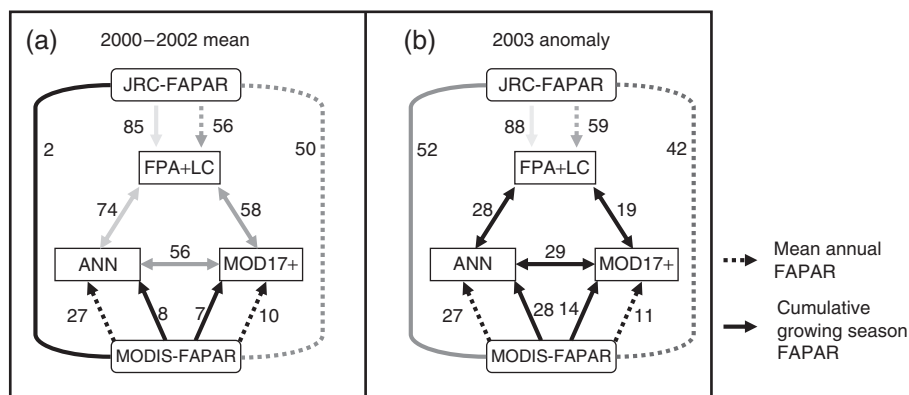
Because the bias from meteorological data is not constant across the European domain, it may also explain why LPJmL, MOD17 + and ANN show lower mean annual GPP in Eastern Europe than FPA + LC. Regional climate models like REMO have difficulties in predicting the climate accurately in south central and the continental Eastern Europe and tend to produce a dry bias (Hagemann et al., 2004; Jacob et al., 2007), which likely propagates to a low biased GPP in this region for the three models. The effect of the NCEP-REMO meteorological data on simulated inter-annual variability of GPP is even more problematic (Jung et al., 2007b) and is likely the cause as to why LPJmL and ANN predict a much stronger 2003 anomaly near the Black Sea than FPA + LC. GPP simulations of the Biome-BGC model forced with NCEP-REMO and ECMWF climatologies were in fact uncorrelated in this region only because of the different meteorological datasets (Jung et al., 2007b).

In the case of LPJmL, the deactivation of crop irrigation in the simulations results in some low

biased GPP of the croplands, in particular towards the east and south of Europe where rainfall declines. The absence of irrigation may also explain why the 2003 anomaly is overemphasized in LPJmL simulations.

The earlier mentioned issues illustrate that the quality of meteorological input data constitutes a serious source of uncertainty for carbon modelling, given that the models are very sensitive to the meteorological input. Quality issues of the meteorological data can explain some differences between FPA + LC and LPJmL, MOD17 + and ANN simulations. However, FPA + LC is still quite similar to MOD17 + and especially ANN regarding the mean GPP pattern, and in the next section, we discuss the possible reasons for this.

Why do the diagnostic models FPA + LC, MOD17 + and ANN tend to converge for the mean GPP pattern but diverge for the 2003 anomaly? The relatively strong inter-correlation of mean spatial GPP fields from different data-oriented models (0.75–0.86; Table 3) indicates an emerging consensus regarding a realistic mean GPP pattern. This makes it reasonable to use mean GPP patterns from diagnostic models as a reference for process-models, the latter display much stronger divergence of the mean GPP pattern of Europe (Jung et al., 2007b). Regarding the spatial pattern of the 2003

Fig. 7 Inter-comparison 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.

anomaly, the situation reverses; the correlations tend 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 3). The obvious question that arises is why are the diagnostic models relatively close to each other regarding the mean GPP pattern but differ substantially regarding the 2003 anomaly pattern?

A close correspondence among the data-oriented model results may be expected, given that all are driven by remote sensing input and linked with CarboEuro flux measurements. We investigated if the similarity of the GPP patterns originates from the remote sensing input by correlating the spatial MODIS and JRC-FAPAR patterns with the corresponding GPP patterns. Fig. 7a presents the correlation analysis for the mean GPP pattern (2000–2002) and shows that (1) the correlation between MODIS-FAPAR and the related MOD17 + and ANN GPP patterns is small (R^2 between 0.07 and 0.27), (2) the correlation of the MODIS and JRC-FAPAR is also relatively small ($R^2 = 0.02$ and $R^2 = 0.5$ for the cumulative growing season and mean annual FAPAR, respectively) and (3) the correlations of the GPP patterns from FPA + LC with MOD17 + and ANN are substantially larger ($R^2 = 0.58$ and $R^2 = 0.74$, respectively). Thus, the convergence of FPA + LC, MOD17 + and ANN regarding the mean GPP pattern does not result from the fact that all use remote sensing input. Instead, the small imprint of the MODIS-FAPAR patterns on MOD17 + and ANN suggests that their GPP patterns are primarily driven by the meteorological input. These findings have important implications: the information contained in

meteorological data is to a large degree redundant for GPP prediction on annual time scale when using the remotely sensed JRC-FAPAR data, given that the meteorology-driven GPP patterns of ANN and MOD17 + are similar to the JRC-FAPAR driven GPP pattern of FPA + LC.

That the spatial GPP pattern of ANN and MOD17 + does not originate primarily from the FAPAR input can also be seen in the correlation analysis for the 2003 anomaly presented in Fig. 7b. The imprint of the MODIS-FAPAR anomaly on GPP patterns from ANN and MOD17 + is again small (R^2 between 0.11 and 0.28). The correlations of the 2003 GPP anomaly patterns among the three diagnostic models are this time also poor (R^2 between 0.19 and 0.29), even between MOD17 + and ANN, which use the same input data. Consequently, the major differences among the three models originate from strong but model-specific sensitivities to the meteorological input of ANN and MOD17 + . Clearly, the model structure of diagnostic models is a substantial source of uncertainty, in particular when studying temporal variations, which are driven by model-specific sensitivities to meteorological conditions in ANN and MOD17 + . Model-specific sensitivity to the meteorological input may partly arise from co-linearity of meteorological variables (e.g. temperature, radiation, vapour pressure deficit), which may cause ambiguous model behaviour in certain circumstances, as the model cannot distinguish their individual effects properly.

The divergence of the three data-oriented GPP models regarding the 2003 heat wave anomaly advises

not to 'blindly' use the patterns of diagnostic models as a benchmark for process models. Model-specific sensitivity to meteorological conditions and in particular capturing drought-related effects is a major issue of diagnostic GPP models that rely on meteorological input. This is in particular true when they do not account for soil moisture but estimate water stress only from atmospheric vapour pressure deficit (Mu *et al.*, 2007). To what extent the remotely sensed JRC-FAPAR picks up inter-annual 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 herbaceous fraction acts as bio-indicator and provides at least the correct direction of change.

To improve our diagnostic GPP models in the future, we have to better understand the relationship between remotely sensed vegetation information and GPP and how these relationships are altered by environmental conditions. In addition, there needs to be clarity of the quality of different alternative remote sensing products because we have noted dramatic differences between the MODIS and JRC-FAPAR datasets (cf. Pinty *et al.*, 2007). Given the current issues of the quality of meteorological data fields, and the model-specific sensitivities to the meteorological input, the simplest approach based only on the JRC-FAPAR seems advantageous.

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. 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 $250 \, \mathrm{g \, C \, m^{-2} \, yr^{-1}}$. Given that herbaceous vegetation together with evergreen forests cover $\sim 80\%$ 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 RUE model MOD17+ and an artificial neuronal network approach on the continental scale, we find that the FAPAR-based GPP estimates show credible patterns of GPP variations. Convergence among the different data-oriented GPP models regarding the mean annual GPP pattern suggests that the mean GPP pattern from diagnostic models may be used for evaluating prognostic models. Divergence among the diagnostic models regarding the pattern of the 2003 anomaly raises caution in using data-oriented models as benchmark for process models regarding inter-annual variations.

This study as well as Rahman *et al.* (2005) and Sims *et al.* (2006b) indicate that just using a simple regression between GPP and a remote sensing product yields more robust results than models that are additionally based on meteorological input. This results from (1) the nature of (some) remotely sensed vegetation properties which sense the response of the vegetation to the environment (i.e. the effect of meteorological variations is largely already implicit) and (2) large uncertainties of meteorological forcing fields and (3) possibly ambiguous model-specific sensitivity to meteorological forcing. A key issue for future improved diagnostic GPP models is to better understand the information that remote sensing products provide in order to best capitalize on them. A well-designed benchmarking exercise for diagnostic GPP models would further help to clarify the performance of different approaches. 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 the availability of ecosystem level, GPP estimates from eddy covariance sites for calibration, it is a 'cheap' and valuable tool to quantify GPP over large regions and a useful dataset for evaluation of biosphere models.

References

Aubinet M, Berbigier P, Bernhofer C *et al.* (2005) Comparing CO_2 storage and advection conditions at night at diferent carboeuroflux sites. *Boundary-Layer Meteorology*, **116**, 63–93.

Beer C, Reichstein M, Ciais P, Farquhar GD, Papale D (2007) Mean annual GPP of Europe derived from its water balance. *Geophysical Research Letters*, **34**, 1–4.

Bondeau A, Smith PC, Zaehle S *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.

Box EO, Holben BN, Kalb V (1989) Accuracy of the AVHRR vegetation index as a predictor of biomass, primary productivity, and net CO_2 flux. *Vegetatio*, **80**, 71–89.

Bradley BA, Jacob RW, Hermance JF, Mustard JF (2007) A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, **106**, 137–145.

Brovkin V, Sitch S, von Bloh W, Claussen M, Bauer E, Cramer W (2004) Role of land cover changes for atmospheric CO_2 increase and climate change during the last 150 years. *Global Change Biology*, **10**, 1253–1266.

Chen Y, Churkina C, Heimann M (2007) *A comparison of regional climate variables between various data sources*. Technical Reports, Max-Planck-Institut für Biogeochemie 8, 36 pp.

Ciais P, Reichstein M, Viovy N *et al.* (2005) Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, **437**, 529–533.

Cook EA, Iverson LR, Graham RL (1989) Estimating forest productivity with thematic mapper and biogeographical data. *Remote Sensing of Environment*, **28**, 131–141.

Desai AR, Richardson AD, Moffat AM *et al.* (2008) Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. *Agricultural and Forest Meteorology*, **148**, 821–838.

Duchemin B, Goubier J, Courrier G (1999) Monitoring phenological key stages and cycle duration of temperate deciduous forest ecosystems with NOAA/AVHRR data. *Remote Sensing of Environment*, **67**, 68–82.

Feser F, Weisse R, von Storch H (2001) Multi-decadal atmospheric modeling for Europe yields multi-purpose data. *EOS Transactions*, **82**, 305–310.

Friedlingstein P, Cox PM, Betts RA *et al.* (2006) Climate-carbon cycle feedback analysis: results from C4MIP model intercomparison. *Journal of Climate*, **19**, 3337–3353.

Gobron N, Pinty B, Aussedat O *et al.* (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**, 1–15.

Gobron N, Pinty B, Aussedat O *et al.* (2008) 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*, **112**, 1871–1883.

Gobron N, Pinty B, Melin F *et al.* (2005) The state vegetation in Europe following the 2003 drought. *International Journal of Remote Sensing Letters*, **26**, 2013–2020.

Gobron N, Pinty B, Verstraete M, 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.

Goward SN, Tucker C, Dye D (1985) North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. *Vegetatio*, **64**, 3–14.

Hagemann S, Machenhauer B, Jones R, Christensen OB, Deque M, Jacob D, Vidale PL (2004) Evaluation of water and energy budgets in regional climate models applied over Europe. *Climate Dynamics*, **23**, 547–567.

Haxeltine A, Prentice IC (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.

Heinsch FA, Zhao MS, Running SW *et al.* (2006) Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 1908–1925.

Jacob D, Bärring L, Christensen OB *et al.* (2007) An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change*, **81**, 31–52.

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

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.

Jung M, Le Maire G, Zaehle S *et al.* (2007a) 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.

Jung M, Vetter M, Herold M *et al.* (2007b) Uncertainties of modeling gross primary productivity over Europe: a systematic study on the effects of using different drivers and terrestrial biosphere models. *Global Biogeochemical Cycles*, **21**, GB4021, doi: 10.1029/2006GB002915.

Kalnay E, Kanamitsu M, Kistler R *et al.* (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77**, 437–471.

Krinner G, Viovy N, de Noblet-Ducoudre N *et al.* (2005) A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system. *Global Biogeochemical Cycles*, **19**, GB1015, doi: 10.1029/2003GB002199.

Litton CM, Raich JW, Ryan MG (2007) Carbon allocation in forest ecosystems. *Global Change Biology*, **13**, 2089–2109.

Lucht W, Prentice IC, Myneni RB *et al.* (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, **296**, 1687–1689.

Luyssaert S, Janssens IA, Sulkava M *et al.* (2007) Photosynthesis drives anomalies in net carbon-exchange of pine forests at different latitudes. *Global Change Biology*, **13**, 2110–2127.

Marcolla B, Cescatti A, Montagnani L, Manca G, Kerschbaumer G, Minerbi S (2005) Importance of advection in the atmospheric CO_2 exchanges of an alpine forest. *Agricultural and Forest Meteorology*, **130**, 193–206.

Moffat AM, Papale D, Reichstein M *et al.* (2007) Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, **147**, 209–232.

Monsi M, Saeki T (1953) Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion. *Japanese Journal of Botany*, **14**, 22–52.

Monteith JL (1965) Light distribution and photosynthesis in field crops. *Annals of Botany*, **29**, 17–37.

Morales P, Sykes MT, Prentice IC *et al*. (2005) Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. *Global Change Biology*, **11**, 2211–2233.

Mu QZ, Zhao MS, Heinsch FA, Liu ML, Tian HQ, Running SW (2007) Evaluating water stress controls on primary production in biogeochemical and remote sensing based models. *Journal of Geophysical Research – Biogeosciences*, **112**, G01012, doi: 10.1029/2006JG000179.

Nakaji T, Ide R, Oguma H, Saigusa N, Fujinuma Y (2007) Utility of spectral vegetation index for estimation of gross CO_2 flux under varied sky conditions. *Remote Sensing of Environment*, **109**, 274–284.

Papale D, Reichstein M, Aubinet M *et al*. (2006) Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, **3**, 571–583.

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

Pinty B, Lavergne T, Vossbeck M *et al*. (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**, D10116, doi: 10.1029/2006JD008105.

Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, 117–134.

Rahman AF, Sims DA, Cordova VD, El-Masri BZ (2005) Potential of MODIS EVI and surface temperature for directly estimating per-pixel ecosystem C fluxes. *Geophysical Research Letters*, **32**, L19404, doi: 10.1029/2005GL024127.

Reichstein M (2006) Integration of FLUXNET and Earth observation data with biogeochemical modelling. *iLEAPS Newsletter*, **3**, 32–34.

Reichstein M, Ciais P, Papale D *et al*. (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**, 634–651.

Reichstein M, Falge E, Baldocchi D *et al*. (2005) On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, **11**, 1424–1439.

Reichstein M, Papale D, Valentini R *et al*. (2007b) Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites. *Geophysical Research Letters*, **34**, L01402, doi: 10.1029/2006GL027880.

Running SW, Nemani RR, Heinsch FA, Zhao MS, Reeves M, Hashimoto H (2004) A continuous satellite-derived measure of global terrestrial primary production. *Bioscience*, **54**, 547–560.

Sakamoto T, Yokozawa M, Toritani H, Shibayama M, Ishitsuka N, Ohno H (2005) A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, **96**, 366–374.

Schaphoff S, Lucht W, Gerten D, Sitch S, Cramer W, Prentice IC (2006) Terrestrial biosphere carbon storage under alternative climate projections. *Climatic Change*, **74**, 97–122.

Sims DA, Luo HY, Hastings S, Oechel WC, Rahman AF, Gamon JA (2006a) Parallel adjustments in vegetation greenness and ecosystem CO_2 exchange in response to drought in a Southern California chaparral ecosystem. *Remote Sensing of Environment*, **103**, 289–303.

Sims DA, Rahman AF, Cordova VD *et al*. (2006b) On the use of MODIS EVI to assess gross primary productivity of North American ecosystems. *Journal of Geophysical Research – Biogeosciences*, **111**, G04015, doi: 10.1029/2006JG000162.

Sitch S, Brovkin V, von Bloh W, van Vuuren D, Assessment B, Ganopolski A (2005) Impacts of future land cover changes on atmospheric CO_2 and climate. *Global Biogeochemical Cycles*, **19**, GB2013, doi: 10.1029/2004GB002311.

Sitch S, Smith B, Prentice IC *et al*. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, **9**, 161–185.

Tedeschi LO (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.

Uppala SM, Kallberg PW, Simmons AJ *et al*. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**, 2961–3012.

van Dijk A, Dolman A, Schulze ED (2005) Radiation, temperature, and leaf area explain ecosystem carbon fluxes in boreal and temperate European forests. *Global Biogeochemical Cycles*, **19**, GB2029, doi: 10.1029/2004GB002417.

Verstraete MM, Gobron N, Aussedat O, Robustelli M, Pinty B, Widlowski JL, Taberner M (2007) An automatic procedure to identify key vegetation phenology events using the JRC-FAPAR products. *Advances in Space Research*, in press.

Vetter M, Churkina G, Jung M *et al*. (2008) Analyzing the causes and spatial pattern of the European 2003 carbon flux anomaly in Europe using seven models. *Biogeosciences*, **5**, 561–583.

White MA, Nemani RR (2006) Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment*, **104**, 43–49.

Xiao XM, Hollinger D, Aber J, Goltz M, Davidson EA, Zhang QY, Moore B (2004) Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sensing of Environment*, **89**, 519–534.

Yang L, Ichii K, White MA *et al*. (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.

Zhang X, Friedl MA, Schaaf CB *et al*. (2003) Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, **84**, 471–475.

Zhao M, Running SW, Nemani RR (2006) Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reana-

lyses. *Journal of Geophysical Research – Biogeosciences*, **111**, G01002, doi: 10.1029/2004JG000004.

Appendix – Alphabetic list of frequently used acronyms

ANN: Artificial Neural Network up-scaling of GPP, TER, and NEP based on flux tower measurements, FAPAR from MODIS, and meteorological data (Papale & Valentini, 2003; Vetter *et al.*, 2008).

APAR: Absorbed Photosynthetic Active Radiation (= FAPAR × PAR) (MJ yr^{-1}).

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

CGS-FAPAR: Cumulative growing season FAPAR value.

DBF: Deciduous Broadleaf Forest.

EBF: Evergreen Needleleaf Forest.

ECMWF: European Centre for Medium-Range Weather Forecasts (refers to the meteorological reanalysis product ERA 40)

ENF: Evergreen Needleleaf Forest.

FAPAR: Fraction of Absorbed Photosynthetic Active Radiation.

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

GPP: Gross Primary Production (g C m^{-2} yr^{-1}).

JRC: Joint Research Centre of the European Union

LPJmL: Lund–Potsdam–Jena managed Land global biosphere model (Sitch *et al.*, 2003; Bondeau *et al.*, 2007).

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; optimized with CarboEurope flux tower data (Reichstein, 2006).

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 (g C m^{-2} yr^{-1}).

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

RUE: Radiation-Use Efficiency (g C MJ^{-1}).

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 dataset produced for terrestrial carbon cycle studies (Jung *et al.*, 2006).

VPD: Vapour Pressure Deficit (Pa).