GRACIELA REDIES FISCHER

OTIMIZAÇÃO MULTI-SÍTIO DE MODELO DE ASSIMILAÇÃO DE CARBONO E APLICAÇÕES EM ECOSSISTEMAS DE CERRADO E PASTAGEM

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Meteorologia Agrícola, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2011

GRACIELA REDIES FISCHER

OTIMIZAÇÃO MULTI-SÍTIO DE MODELO DE ASSIMILAÇÃO DE CARBONO E APLICAÇÕES EM ECOSSISTEMAS DE CERRADO E PASTAGEM

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Meteorologia Agrícola, para obtenção do título de *Doctor Scientiae*.

APROVADA: 29 de julho de 2011.

Paulo José Hamakawa

Edson Luís Nunes

Ana Cláudia Mendes Malhado

Regina Célia dos Santos Alvalá
(Coorientadora)

Marcos Heil Costa
(Orientador)

A única arma eficaz contra as idéias são idéias melhores.
(Alfred Whitney Griswold)

Otimismo significa esperar o melhor, mas confiança significa saber lidar com o pior.
(Max Gunther)

O bom humor espalha mais felicidade que todas as riquezas do mundo. Vem do hábito de olhar para as coisas com esperança e de esperar o melhor e não o pior.
(Alfred Montapert)

Há duas fontes perenes de alegria pura: o bem realizado e o dever cumprido.
(Eduardo Girão)

O homem sábio é aquele que não se entristece com as coisas que não tem, mas se rejubila com as que têm.
(Epicteto)

À minha mãe Elma Redies Fischer.

Ao meu pai Elmo Fischer.

Ao meu companheiro Leonardo José Gonçalves Aguiar.

AGRADECIMENTOS

- À Deus, por dar-me forças, saúde e sabedoria e também por ter sempre iluminado meu caminho.
- Aos meus pais Elmo e Elma Fischer, pela educação, pelo carinho, incentivo e apoio incondicional em todos os momentos.
- Ao meu grande companheiro, Leonardo José Gonçalves Aguiar, pelas contribuições, pelo apoio e incentivo nas horas difíceis, pelo amor, a amizade, força e dedicação.
- À Universidade Federal de Viçosa (UFV) e ao Departamento de Engenharia Agrícola, pela oportunidade de realizar o curso de pós-graduação em Meteorologia Agrícola.
- À Coordenadoria de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), pela concessão de bolsa de estudo.
- Ao meu orientador, Professor Marcos Heil Costa, pelos ensinamentos, atenção, apoio profissional e pelas sugestões dadas durante a elaboração deste trabalho.
- À co-orientadora Regina Célia dos Santos Alvalá pelas sugestões e contribuições para este trabalho.
- Aos professores do programa de Pós-Graduação pelo conhecimento transmitido.
- Ao Fabrício, pela amizade e pela grande e valiosa ajuda com os scripts para calibração do modelo.
- À Gabrielle, pela amizade e pela ajuda nos scripts NCL.

- Aos amigos Ana e Richard, pelo apoio em todos os momentos, pelas contribuições ao longo deste trabalho. Pelos jantares, almoços, lanchinhos, conversas, cantorias, aventuras, amizade e momentos divertidos juntos com os cachorros (Malhadinha e Pitu).
- Aos amigos Luciana e ao Luiz, pela amizade, conversas, pelos churrascos e momentos de descontração (Guri e Pitu).
- Ao Edson, pela grande ajuda com o modelo RATE.
- Aos colegas do Grupo de Pesquisa, Francisca, Leydimere, Hewlley, Patrícia, Lívia, Paulinha, Marcos Paulo, Vitor, Christiane, Varejão, Alessandro, João Paulo, pelo companheirismo, conversas, lanchinhos e pela boa convivência.
- À Letícia Braida, pela amizade e carinho.
- À Graça, excelente secretária da Meteorologia Agrícola, pela ajuda e amizade ao longo do doutorado.
- Aos amigos da Meteorologia pela amizade e convívio. Em especial os amigos Evandro, Leonardo Neves (Coelho), Maria Emília e Raniéri.
- À amiga Sabrina Rodth, pela amizade e pela acolhida quando cheguei a Viçosa.
- A todos da minha família, pelo carinho, apoio e incentivo.
- Às amigas do pilates, Rayana e Cristiana, pela amizade e confiança de que tudo daria certo.
- À minha sogra e aos meus cunhados, pela amizade, apoio e incentivo.
- A todos que contribuíram direta ou indiretamente para a realização deste trabalho.

BIOGRAFIA

GRACIELA REDIES FISCHER, filha de Elmo Fischer e Elma Redies Fischer, nasceu no dia 05 de maio de 1983, na cidade de Pelotas – RS.

Iniciou a graduação em Meteorologia em março de 2001, obtendo o título de Bacharel em Meteorologia, pela Universidade Federal de Pelotas (UFPel), em fevereiro de 2005.

Em fevereiro de 2007, concluiu o curso de pós-graduação, em nível de Mestrado, em Meteorologia na Universidade Federal de Pelotas (UFPel).

Em março de 2007 iniciou o curso de pós-graduação, em nível de Doutorado, em Meteorologia Agrícola na Universidade Federal de Viçosa (UFV), na área de micrometeorologia de ecossistemas.

SUMÁRIO

LISTA DE ABREVIATURAS

APAR	Absorbed photosynthetically active radiation
BAN	Bananal site
CASA	Carnegie-Ames-Stanford approach Biosphere Model
DW	Multi-site calibration weighting the data according to the number of data points available for each site
E_{max}	Maximum bias error
FAPAR	Fraction absorbed of photosynthetic active radiation
FNS	Fazenda Nossa Senhora site
GPP	Gross primary production
HDF-EOS	Hierarchical data format developed by National Center for Supercomputer Application
IBIS	Integrated Biosphere Simulator
K77	km 77 site
LAI	Leaf area index
LBA	Large-Scale Biosphere-Atmosphere Experiment in Amazonia
LBA-DMIP	LBA Model Intercomparison Project
LSM	Land Surface Model
LSX	Land-surface transfer scheme model
MAE	Mean absolute error
MBE	Mean bias error
MODIS	Moderate Resolution Imaging Spectroradiometer
MOD12Q1	Land cover MODIS product
MOD15A2	LAI and FAPAR MODIS product
MOD17A3	NPP MODIS product
NCAR	National Center for Atmospheric Research

NCEP	National Centers for Environmental Prediction
NCL	NCAR Command Language
netCDF	network Common Data Form format
NEE	Net ecosystem exchange
NPP	Net primary production
OM	Optimization Method
PAR	Photosynthetic active radiation
PDG	Pé de Gigante site
PFT	Plant Functional Type
r	Correlation coefficient
r^2	Determination coefficient
RATE	Regional Algorithm for monitoring the carbon assimilation of Terrestrial Ecosystems
RMSE	Root mean square error
SiB2	Simple Biosphere Model
SITE	Simple Tropical Ecosystem Model
SW	Multi-site calibration using the same weight for each site

LISTA DE SÍMBOLOS

A_g	Gross photosynthesis
A_n	Net photosynthesis rate
awc	Available water content of the soil
α_4	Intrinsic quantum efficiency for C_4 plants
b	Linear relationship intercept
$CO2_i$	CO_2 concentration in the intercellular spaces of the leaves
C_s	Leaf boundary layer CO_2 concentration
d	Index of agreement
D_f	Carbon mass of dead fine roots
f	Represents the adjustment measures
f_d	Soil temperature function in the d layers
f_g	Soil temperature function in the g layers
$F(\theta)$	Objective function for four optimization methods
γ	Leaf respiration coefficient
Γ	Compensation point for gross photosynthesis
g_d	Soil moisture function in the d layers
g_g	Soil moisture function in the g layers
G_s	Canopy stomatal conductance
h_u	Respiration rate of leaves litter
h_f	Respiration rate of fine roots litter
J_c	Gross photosynthesis rate limited by the activity of the Rubisco enzyme
J_e	Gross photosynthesis rate limited by light
K_0	Michaelis constant
K_c	Michaelis constant

K_f	Multiplier of respiration rate of fine roots
K_u	Multiplier of respiration rate of leaves litter
L_u	Carbon mass of leaf litter
m	Stomatal conductance angular coefficient
n	Number of data points available for each site
O	Observed data
PSN_{net}	Net photosynthesis in daily basis
R_a	Autotrophic respiration
R_f	Respiration rate of the fine roots
R_g	Annual growth respiration
R_H	Heterotrophic respiration
r_h	Relative humidity
R_{lr}	Daily maintenance respiration of leaves and fine roots
R_m	Maintenance respiration of live cells in woody tissues
R_r	Respiration rate of the thick roots
R_s	Respiration rate of stems
R_u	Respiration rate of leaves
S	Simulated data
S_t	Soil moisture stress
St_m	Soil moisture stress coefficient
T_f	Function of minimum air temperature
θ	Set of model's parameters to be optimized
Θ	Constrain set of model's parameters to be optimized
V_m	Maximum Rubisco enzyme capacity
V_{max}	Maximum rubisco activity at 15ºC
VPD_f	Function of vapor pressure deficit

LISTA DE FIGURAS

Página

LISTA DE TABELAS

Página

RESUMO

FISCHER, Graciela Redies, D.Sc., Universidade Federal de Viçosa, julho de 2011. **Otimização multi-sítio de modelo de assimilação de carbono e aplicações em ecossistemas de cerrado e pastagem.** Orientador: Marcos Heil Costa. Coorientadora: Regina Célia dos Santos Alvalá.

Diante de cenários de mudanças climáticas globais, estudos usando modelos de superfície terrestre são necessários para avaliar o impacto do aumento do CO_2 e o balanço de carbono em ecossistemas terrestres. Entretanto, uma boa calibração desses modelos é importante para que esses modelos representem da melhor maneira possível os processos do ecossistema. Geralmente, os modelos são calibrados para um sítio e, então, os mesmos parâmetros calibrados são aplicados para outros sítios com o mesmo tipo funcional de planta. Tendo em vista o exposto, este trabalho possui dois objetivos principais: Inicialmente, realizar uma calibração multi-sítio, utilizando dois sítios simultaneamente, e avaliar seu desempenho; E, então, fazer a aplicação dos resultados da calibração multi-sítio para gerar estimativas de produção primária líquida (NPP, em inglês) em ecossistemas de cerrado e pastagem na América do Sul. Para investigar a calibração multi-sítio, foi realizada a calibração da troca líquida do ecossistema (NEE, em inglês) para dois sítios de pastagem na Amazônia, analisando 20 diferentes funções objetivo, sendo cinco medidas de ajuste sujeitas a quatro métodos de otimização – duas otimizações individuais (uma para cada sítio) e duas otimizações multi-sítio (uma com o mesmo peso para cada sítio e a outra com peso proporcional ao tamanho da série de dados). As cinco medidas de ajuste utilizadas foram: coeficiente de correlação (r), raiz do erro quadrado médio (RMSE), erro médio absoluto (MAE), erro de viés médio (MBE) e erro de viés máximo (E_{max}). Os resultados da calibração do modelo indicam que com algumas

restrições da escolha da função objetivo, a calibração multi-sítio é possível e produz resultados consistentes entre os sítios. A escolha da função objetivo deve ser baseada na intenção de uso do modelo. Para modelos de simulações de curto prazo é recomendado o uso do método de otimização com o mesmo peso para cada sítio, usando o MAE como medida de ajuste. Por outro lado, para modelos de simulações de longo prazo, o método indicado é o mesmo peso para cada sítio, utilizando o E_{max} como função objetivo. Visto que neste trabalho a prioridade são simulações de longo prazo, aplicou-se o método do peso para cada sítio utilizando o E_{max} como função objetivo para gerar estimativas de NPP em ecossistemas de cerrado e pastagem para a América do Sul. Os parâmetros calibrados foram aplicados no Algoritmo RATE (baseado no modelo SITE), que é um algoritmo de monitoramento regional da taxa de fixação de carbono usando dados de sensoriamento remoto derivados dos produtos do MODIS (cobertura do solo, índice de área foliar e fração absorvida da radiação fotossinteticamente ativa). Os resultados indicam uma pronunciada variabilidade sazonal nas estimativas mensais de NPP para o cerrado e para a pastagem no ano de 2010, mostrando também uma conexão da variabilidade de NPP com a variabilidade climática da região – períodos de alta e baixa precipitação e temperatura controlando diferenças sazonais na fixação de carbono. A variabilidade interanual de NPP foi maior para a pastagem do que para o cerrado, talvez indicando maior suscetibilidade da pastagem à variabilidade climática. Os resultados apresentaram uma clara diferença entre as estimativas de NPP do RATE e do MODIS, indicando que mais estudos são necessários para validar estratégias alternativas de estimar NPP.

ABSTRACT

FISCHER, Graciela Redies, D.Sc., Universidade Federal de Viçosa, July, 2011. **Multi-site otimization of carbon assimilation model and applications in cerrado and grassland ecosystems.** Adviser: Marcos Heil Costa. Co-Adviser: Regina Célia dos Santos Alvalá.

Front of global climate change scenarios, studies using land surface models are needed to assess the impact of increased CO_2 and the carbon balance in terrestrial ecosystems. However, to be effective these models must be carefully calibrated to accurately represent ecosystem processes. Generally, the model is calibrated for one site and then used with the same set of calibrated parameters, either in other sites or in a whole region with the same plant functional type. In view of the exposed, this research has two main objectives: first, make a multi-site calibration, using two sites simultaneously, and evaluate their performance; and then, make application of the results of multi-site calibration to generate estimates of net primary production (NPP) in cerrado and grassland ecosystems in South America. To investigate the multi-site calibration, the calibration was performed to net ecosystem exchange (NEE) for two pasture sites in the Amazon. Twenty different objective functions, being five adjustment measures subject to four optimization methods (two individual optimizations - one for each site - and two multi-site optimizations - one with the same weight to each site and the other with a weight proportional to the size of the dataset) are evaluated to investigate the consistency of the results in a multi-site model optimization. The five adjustment measures were used: correlation coefficient (r), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE) and maximum bias error (E_{max}). Our results indicate that, with some

restrictions regarding the choice of objective function, multi-site calibration is possible and produces consistent results across sites. We recommend that the choice of objective function should be based on the intended use of the model. On a short time scale we recommend the site-weighted method using mean absolute error as adjustment measure. For longer times scales we recommend using the site-weighted maximum bias error. For this research the priority is long-term simulations, so we applied the site-weighted method using maximum bias error to generate estimates of NPP in cerrado and grassland in South America. The calibrated parameters were applied in Algorithm RATE (based on the SITE vegetation model), which is an algorithm for regional monitoring of the carbon fixation rate using land cover, LAI and FAPAR datasets derived from a remote sensing satellite (MODIS) product. The results indicate a pronounced seasonal variability in monthly NPP estimates for cerrado and grassland in 2010 linked to the climate variability of the region - periods of high and low precipitation and temperature driving seasonal differences in carbon fixation. Interannual NPP variability was higher for grassland than cerrado, perhaps indicating that grassland is more susceptible to climate variability. There were also clear differences between MODIS NPP estimates and those of RATE, indicating that further studies are urgently needed to rigorously validate these alternative strategies to estimate NPP.

GENERAL INTRODUCTION

Several research projects have been conducted in order to better understand carbon cycle dynamics under anthropogenic climate change (Grant et al., 2009; Malhi et al., 2009). The study of carbon balance in terrestrial ecosystems is very relevant to assess how ecosystems respond to increasing CO_2 levels and if ecosystems are absorbing CO_2 from the atmosphere or emitting it. Thus, net primary production (NPP) is an important aspect of studies of carbon balance due to its direct connection to the carbon cycle. NPP is a measure of the carbon sequestration rate by the ecosystem and can be affected by climatic conditions and others factors.

Field measurements of NPP are technically difficult and require a considerable investment of time and human resources. Moreover, these measurements are insufficient to fully evaluate the spatio-temporal variability of NPP (Nayak et al., 2010). Thus, regional estimates of NPP obtained through other methodologies are essential to improve knowledge of the spatio-temporal variability of NPP and the magnitude of carbon storage by ecosystems. Perhaps understandably, ecosystems with apparently low potential for carbon assimilation, such as the cerrado or grassland ecosystems, have received relatively little attention in NPP studies and most knowledge of carbon dynamics in these ecosystems in Brazil is derived from global studies.

Land surface models are probably the major tool to perform regional estimates of NPP. Frequently, these models combine remote sensing and climate information to improve simulations. Specifically, remote sensing products from Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the Terra and Aqua satellites are often used as input data and to compare with the output of the land surface model simulations (Potter et al., 2009; Nayak et al., 2010; Yuan et al., 2010).

Models that are used to estimate NPP and other variables should be calibrated to provide accurate and realistic simulations of atmosphere-vegetation-soil processes (Liu et al., 2003). To generate the most robust regional estimates the parameters that represent these processes within the model should, ideally, be specifically calibrated for each ecosystem (Groenendijk et al., 2011). In other words, the calibration of the model will largely define the performance of the model.

Generally, land surface models are calibrated for a single site of one particular vegetation class or plant functional type (PFT). This methodology is often applied because there are insufficient empirical data available; however, more data are becoming available allowing the development of different methodologies. One such possibility is multi-site calibration, a method that it is frequently used in studies of the hydrology, but which has not been frequently applied or tested in land surface and ecosystem modelling. Multi-site calibration consists of calibrating two or more sites in the same vegetation class simultaneously.

In view of the potential importance of improved regional estimates of NPP that is, in turn, dependent on the accurate calibration of the estimation models, the objective of this thesis is evaluate the potential for multi-site calibration and to apply this methodology to obtain NPP estimates for cerrado and pasture ecosystems in

South America. The thesis is organized in two chapters: Chapter 1 investigates the performance of multi-site calibration for models of net ecosystem exchange (NEE) for two pasture sites in Amazonia using a simple terrestrial ecosystem model (SITE – Santos and Costa, 2004). Chapter 2 describes the implementation of multi-site calibration to perform estimates spatio-temporal patterns of NPP over grassland and cerrado using the RATE Algorithm (Nunes, 2008).

CHAPTER 1

MULTI-SITE LAND SURFACE MODEL OPTIMIZATION – AN EXPLORATION OF OBJECTIVE FUNCTIONS

1.1. INTRODUCTION

Land surface models are an essential tool in climate modeling, and coupled climate-vegetation models have been widely used in investigations about the effects of changes in land use, land surface processes, carbon cycle and its implications for the Earth's climate (Foley et al., 2000; McGuire et al., 2001; Williams et al., 2009; Costa and Pires, 2010). Understanding land use changes in the context of climate modelling is important because coupled climate-biosphere models can be used in various scenarios of land use change. However, to provide realistic results and forecasts, models must be calibrated to accurately represent ecosystem processes (Liu et al., 2003).

In terrestrial ecosystem models several processes are represented by parameters which need to be specifically calibrated for each ecosystem – these parameters defining the performance of the model (Groenendijk et al., 2011). Given the frequent lack or inability to obtain direct measurements, some of these parameters must be obtained through optimization methods in order to minimize the differences between observed data and the model outputs. Several optimization methods have been used for parameter estimation in ecosystems models, such as genetic algorithms (D'heygere et al., 2006), gradient methods (Wang et al., 2001; Rayner et al., 2005), Kalman filters (Mo et al. 2008; Ju et al, 2009; Zhu et al. 2009)

and global search methods (Braswell et al., 2005; Knorr and Kattge, 2005). However, it is the choice of adjustment measure rather than the choice of optimization method that has the greatest impact on the model results (Trudinger et al., 2007).

Various adjustment measures have been proposed to identify and account the errors between simulated and observed datasets. The most commonly used measures are the correlation coefficient (r), mean absolute error (MAE), root mean square error (RMSE), mean bias error (MBE), and slope of the least-squares regression between simulated and observed data (Willmott, 1982; Legates and McCabe, 1999; Willmott and Matsuura, 2005).

Another important issue in model optimization is that models are generally calibrated for a single site that is characterized by a particular vegetation class or plant functional type (PFT). After the calibration, the model is applied to other sites or to an entire region with the same PFT. The calibrated parameter set, however, may not be representative for other sites or for the wider region (Groenendijk et al., 2011). Alternative solutions that perform multi-site calibration (simultaneous calibration at two or more sites) are therefore desirable.

Multi-site calibration has recently become more common due to greater data availability, improved model sophistication and better computing capabilities (White and Chaubey, 2005). Indeed, several different approaches to multi-site model calibration have been considered (White and Chaubey, 2005; Cao et al., 2006; Bekele and Nicklow, 2007; Zhang et al., 2008; Zhang et al., 2010), although most examples come from the hydrology literature. In one of the earliest studies, White and Chaubey (2005) applied multi-site calibration to a watershed model, using an objective function that simultaneously minimized errors to several variables and sites. They concluded that multi-site calibration was able to more closely predict

measured values and could be used to accurately predict watershed response for various outputs.

Bekele and Nicklow (2007) considered two calibration methods for a watershed model; the first method used specific objective functions to fit different portions of the time series. In the second method the calibration was performed by simultaneously using data from multiple gauging stations. The results indicated that the second calibration method outperformed the first. Zhang et al. (2008) attempted to optimize 16 parameters of a watershed model through the application of a single objective optimization method and a multi-objective optimization algorithm to a single site and to three sites concurrently. They concluded that parameters estimated by optimizing the objective function at three sites consistently produced better goodness-of-fit than those obtained by optimization at a single site.

In contrast to hydrological studies, multi-site optimization of land surface ecosystem models is still in an early stage of development and many questions remain unanswered. It is still unclear, for instance, whether model parameters should have the same optimized value at different sites with the same PFT. Some parameters may be strongly species-dependent and their true value may vary according to the composition at each site, even for a same PFT. This effect could potentially be isolated by using different sites covered by the same single species – a common situation for agricultural land uses. However, even if isolation of ecosystem composition is possible, other questions arise. One fundamental unanswered

question, and the subject of this chapter, is how model parameters will vary according to the objective function used during multi-site optimization.

To investigate this issue we use a simple terrestrial ecosystem model (SITE – Santos and Costa, 2004) to perform a multi-site calibration of net ecosystem exchange (NEE) for two pasture sites in Amazonia. Model performance is optimized using 20 different objective functions: Five adjustment measures subject to four optimization methods – two single-site optimizations (one per site) and two multi-site optimizations (varying on the weight given to each site: the same weight or a weight proportional to the duration of the time series in each site).

1.2. MATERIAL AND METHODS

1.2.1. Experimental Sites

The calibration was performed for two experimental pasture sites located in the Amazon. The sites are part of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) towers network.

The first pasture site is located at the Fazenda Nossa Senhora ranch (hereafter referred to as site FNS) (10°45'S, 62°21'W, 230 m), in Ouro Preto d'Oeste, Rondônia, Brazil. This site is in the center of a deforested area with an approximate radius of 50 km – deforestation was caused by a fire in 1977 to clear land for crop cultivation. Since the early 1980s the area has been uniformly covered by the grass *Brachiaria brizantha*. Hodnett et al. (1996) classifies the soil as a medium textured red-yellow podzol. The soil is deep to reach the rocky and has a sandy texture on the surface, increasing the percentage of clay with depth. The FNS soil has 80% of sand, 10% of silt and 10% of clay.

The second site is on a farm 77 km along the Santarém-Cuiabá highway (hereafter referred to as site K77) (03°01'S, 54°53'W), near Santarém, Pará, Brazil. The forest was cleared in 1990 and the field was planted with the same grass species as the first site, *Brachiaria brizantha*. In November 2001 the site was burned and plowed for rice cultivation. The K77 soil is classified as yellow latosol and has 18% of sand, 02% of silt and 80% of clay.

1.2.2. Model Description

SITE (Simple Tropical Ecosystem Model) is a simplified dynamic vegetation model of tropical ecosystems developed by Santos and Costa (2004). This model is based on previously developed models, mainly LSX (Pollard and Thompson, 1995), LSM (Bonan, 1996), IBIS (Foley et al., 1996) and SiB2 (Sellers et al., 1996) and provides a simple simulation of the fluxes of CO_2, water and energy, as well as the dynamics of carbon in the ecosystem. It operates through modelling the relationships between several fundamental ecosystem processes: canopy infrared radiation balance, solar radiation balance, aerodynamic processes, canopy physiology and transpiration, balance of water intercepted by the canopy, transport of mass and energy in the atmosphere, soil heat flux, soil water flux and carbon balance. Although SITE is an intentionally simple model, it has the necessary complexity to represent the main processes responsible for ecosystem functioning.

SITE is a dynamical punctual model that uses an integration time step (*dt*) of one hour. The model is structured with one layer of canopy and two layers of soil, the first layer is in 0.10 cm of the depth and the second layer with the depth of the 5 m. The main output variables of the model are latent heat flux, sensible heat flux,

water vapor flux and net ecosystem exchange (NEE). In this study, we only optimize for NEE.

A full specification of the SITE model is provided by Santos and Costa (2004). Here, we present a brief description of the part of the model that is relevant to understand the calibration of parameters and the optimized output. For convenience, the calibrated parameters are identified by a superscript asterisk (*) and are summarized in Table 1.1.

NEE is expressed as the difference between soil heterotrophic respiration (R_H) and the net primary production (NPP). Negative values of NEE indicate assimilation of carbon by the ecosystem and positive values indicate carbon release by the ecosystem.

$$NEE = R_H - NPP \tag{1}$$

Net primary production (NPP, kg-C m^{-2} s^{-1}) is expressed as a function of gross photosynthesis (A_g, mol-CO_2 m^{-2} s^{-1}) and canopy autotrophic respiration (leaves, roots).

$$NPP = (1-\eta)\int (A_g - R_u - R_f)dt \tag{2}$$

where η = 0.3 is the fraction of carbon lost due to growth; A_g (Eq. 3) is the minimum between J_e (gross photosynthesis rate limited by light) and J_c (gross photosynthesis rate limited by the activity of the Rubisco enzyme); R_u is the respiration rate of leaves (mol-CO_2 m^{-2} s^{-1}) and R_f is the respiration rate of the fine roots.

$$A_g = \min(J_e, J_c) \tag{3}$$

J_e is expressed by:

$$J_e = \alpha_4 APAR\left(\frac{CO_{2i} - \Gamma}{CO_{2i} + 2\Gamma}\right) \tag{4}$$

where α_4 is the intrinsic quantum efficiency for C_4 plants (mol CO_2 mol^{-1} photons); *APAR* is the density flux of the absorbed photosynthetically active radiation (mol photons m^{-2} s^{-1}); CO_{2i} is the CO_2 concentration in the intercellular spaces of the leaves; and $\Gamma = [O_2]/2\tau$ is the compensation point for gross photosynthesis (mol mol^{-1}); O_2 is the oxygen concentration and τ is the CO_2/O_2 ratio of kinetic parameters.

J_c is given by:

$$J_c = \frac{V_m(CO_{2i} - \Gamma)}{CO_{2i} + K_c(1 + ([O_2]/K_0))} \tag{5}$$

where K_c and K_0 are two levels of the Michaelis constant: K_c is used to fix CO_2 and K_0 to inhibit oxygenation (mol mol^{-1}), where $K_c = 1.5x10^{-4}$ and $K_0 = 0.25$ a (15°C).

$$CO_{2i} = C_S - \frac{1.6A_n}{G_S} \tag{6}$$

where C_s is the leaf boundary layer CO_2 concentration; A_n is the net photosynthesis rate; and G_S is canopy stomatal conductance.

$$C_S = CO_{2a} - \frac{A_n}{G_b} \tag{7}$$

where CO_{2a} is initial CO_2 concentration and G_b is the boundary layer conductance for CO_2.

$$G_S = m^* \frac{A_n r_h}{C_S} + b \tag{8}$$

where m^* is the stomatal conductance coefficient; r_h is the relative humidity; b is the linear relationship intercept.

$$A_n = A_g - R_u \tag{9}$$

where A_g is the gross photosynthesis R_u is the respiration rate of leaves (mol-CO_2 m^{-2} s^{-1})

R_u is calculated as:

$$R_u = \gamma^* V_m \tag{10}$$

where γ^* is the leaf respiration coefficient of Rubisco enzyme. V_m is the maximum Rubisco enzyme capacity (V_m, μmol-CO_2 m^{-2} s^{-1}) for the carboxylase function calculated through the V_{max}^* parameter (at 15ºC) and modified by soil moisture stress (St) and temperature stress factor (T_{vm}) functions.

$$V_m = V_{\max}^* T_{vm} St \tag{11}$$

$$T_{vm} = \frac{\exp[3500((1/288.16) - (1/T_f))]}{[1+\exp(2 - 0.40T_f)][1+\exp(0.40T_f - 20)]} \tag{12}$$

where T_f is leaf temperature (K). $T_f = T_u - 273.16$, where T_u is upper canopy leaf temperature (K).

$$St = \frac{1-\exp\left(St_m^* awc\right)}{1-\exp\left(St_m^*\right)} \tag{13}$$

where St_m^* is the parameter for soil moisture stress and awc is the available water content of the soil.

Soil heterotrophic respiration (R_H) is the release of CO_2 to the atmosphere during process of decomposition of organic matter. R_H is based on the decomposition of leaves and fine roots by bacteria, and is a function of soil temperature, soil moisture and carbon mass in the litter. R_H is calculated by sum of the products of the temperature and moisture function, and the decomposition rate reservoir of soil organic matter.

$$R_H = f_g g_g K_u^* h_u L_u + \frac{(f_g + f_d)}{2}\frac{(g_g + g_d)}{2} K_f^* h_f D_f \tag{14}$$

where f_g and f_d represent soil temperature functions in the g and d soil layers (dimensionless); g_g and g_d are soil moisture functions in the g and d soil layers (dimensionless); h_u is the respiration rate of leaves litter (kg-C kg^{-1}-C s^{-1}); h_f is the

respiration rate of fine roots litter (kg-C kg^{-1}-C s^{-1}); K_u^* and K_f^* are calibration constants that multiply h_u and h_f, respectively; L_u is the carbon mass of leaf litter; D_f is the carbon mass of dead fine roots (kg-C m^{-2}).

1.2.3. Input and Flux Data

The model is forced by hourly meteorological variables, such as air temperature, specific humidity, wind speed, incident short wave radiation, incident long wave radiation and precipitation. These data were recorded from January 1999 to December 2001 in the FNS site and from January to December 2001 in the K77 site. Hourly carbon flux data were collected from February 1999 to December 2001 (a total of 6264 data points) at FNS and from January to November 2001 (a total of 2768 data points) at K77. Our analysis of the K77 site is restricted to the above time periods because the site was burnt in November 2001 and subsequently converted to rice cultivation.

In order to avoid the excessive weight attributable to respiration parameters during the night, we use only data collected between 07.00 and 17.00 hours for NEE optimization. Moreover, we filtered the carbon flux data to eliminate clearly spurious values - for example values of carbon flux lower than -20 kg-C ha^{-1} h^{-1}.

1.2.4. Optimization Method (OM)

We use the global search method for model optimization. This method searches for the best solution in global parameters space. As a consequence of a lack

of knowledge about the kind of parameter interactions could happen in a multi-site optimization – and that could potentially lead to unexpected local minima – we decided to search for every possible combination of calibrated parameters in each parameter range.

Based on previous sensitivity analysis (by Imbuzeiro (2005) and Nunes (2008)) we chose to calibrate only the parameters most likely to influence the result of the model and to respond to changes in the sites to be calibrated: leaf respiration coefficient (γ^*, dimensionless), stomatal conductance angular coefficient (m^*, dimensionless), maximum rubisco capacity ($V_{max}{}^*$, μmol-CO_2 m^{-2} s^{-1}), soil moisture stress coefficient ($St_m{}^*$, dimensionless), $K_u{}^*$, the multiplier of respiration rate of leaves litter (h_u, kg-C kg^{-1}-C s^{-1}), and $K_f{}^*$, the multiplier of respiration rate of fine roots litter (h_f, kg-C kg^{-1}-C s^{-1}). We then wrote a script to run the model, testing each parameter against the pre-defined appropriate range of values (Table 1.1). We performed 2,347,380 simulations to test the six model parameters, separately analyzing each possible combination.

Table 1.1. Parameters calibrated and range of values tested.

Calibrated parameters	**Range of values tested**
γ	0.010 – 0.050 (changing 0.005, dimensionless)
m	1.0 – 10.0 (changing 1.0, dimensionless)
V_{max}	10.0 – 120.0 (changing 5.0, μmol-CO_2 m^{-2} s^{-1})
St_m	-1.0 – -7.0 (changing 1.0, dimensionless)
K_u	1.0 – 5.0 (changing 0.5, dimensionless)
K_f	1.0 – 5.0 (changing 0.5, dimensionless)

To initialize the model we use the following values: carbon stored in leaves (0.360 kg-C m^{-2}) and fine roots (0.100 kg-C m^{-2}); initial fraction of soil moisture (0.21, in both layers); initial mass litter from leaves (0.10 kg-C m^{-2}) and fine roots (0.09 kg-C m^{-2}); and volume fraction of organic matter (0.547).

To spin-up the model, the available input data were replicated four times. For example, for the FNS site three years of input data were available, so the model was run for 12 years – although only the last three years of data were used to analyze the results.

Four optimization methods were used to optimize the model: individual site calibration (two single-site optimizations – one per site) and two types of multi-site calibration – the first (SW) using the same weight for each site, and the second (DW) weighting the data according to the number of data points available for each site (n_{FNS}= 6264 and n_{K77}= 2768).

The general optimization problem was to find the set of parameters θ that minimizes a given objective function F:

$$\min_{\theta \in \Theta} F_{ind}(\theta) = f_{ind}(\theta) \tag{15}$$

$$\min_{\theta \in \Theta} F_{DW}(\theta) = \frac{n_{FNS} f_{FNS}(\theta) + n_{K77} f_{K77}(\theta)}{n_{FNS} + n_{K77}} \tag{16}$$

$$\min_{\theta \in \Theta} F_{SW}(\theta) = \frac{f_{FNS}(\theta) + f_{K77}(\theta)}{2} \tag{17}$$

where F(θ) is the objective function for four optimization methods (individual – for two sites, DW and SW), θ is the set of model's parameters to be optimized, Θ is the constrain set, and *f* represents the adjustment measures (Fu et al., 2005). To assess each OM we used a range of different adjustment measures (as discussed in *Section 1.2.5* below).

1.2.5. Adjustment Measures

Five adjustment measures were calculated to evaluate the objective functions: correlation coefficient (*r*), root mean square error (RMSE), mean absolute

error (MAE), mean bias error (MBE) and maximum bias error (E_{max}). Some of these measures estimate the dispersion of the data (e.g. *r*, RMSE and MAE), while MBE and E_{max} assess the long-term adjustment of the model. While optimizing these adjustment measures, we maximize *r* and minimize all the others. A definition and more detailed description of the properties of each adjustment measure follow:

The correlation coefficient (*r*) is a metric of the degree of linear relationship between two quantitative variables and can be used to indicate the precision of the model. The coefficient is obtained by dividing the covariance of two variables by the product of their standard deviations (Eq. 18). However, this measure is not particularly robust and is insensitive to additive or proportional differences between the observed data (O) and the simulated data (S).

$$r = \left[\frac{\sum_{i=1}^{n} \left(S_i - \overline{S_i}\right)\left(O_i - \overline{O_i}\right)}{\sqrt{\sum_{i=1}^{n} \left(S_i - \overline{S_i}\right)^2 \sum_{i=1}^{n} \left(O_i - \overline{O_i}\right)^2}} \right] \tag{18}$$

The root mean squared error (RMSE) is the square root of the average squared distance of a data point from the fitted line (Eq. 19). Willmott and Matsuura (2005) state that RMSE is sensitive to large errors, and tends to give stronger emphasis on fitting of peak output values. This means that it is very useful when large errors are particularly undesirable. RMSE expresses errors in the units and scale of the variables of interest.

$$RMSE = \left[\frac{\sum_{i=1}^{n} \left(S_i - O_i\right)^2}{n} \right]^{1/2} \tag{19}$$

The mean absolute error (MAE) is the average absolute error of differences between simulated and observed data (Eq. 20). MAE is very similar to RMSE, but is less sensitive to large forecast errors (Willmott, 1982) and is preferred for small or limited datasets. MAE expresses errors in the units and scale of the variables of

interest. These three metrics only represent the dispersion of the data and do not provide any information about the variability of the errors or provide a metric of long-term errors.

$$MAE = \frac{\sum_{i=1}^{n} |S_i - O_i|}{n} \tag{20}$$

The mean bias error (MBE) is the mean difference between simulated values by model and the observed data (Eq. 21). It is usually intended to indicate average model 'bias', that is, average overestimation or underestimation (Willmott and Matsuura, 2005). MBE should be interpreted cautiously since it is inconsistently related to typical-error magnitude.

$$MBE = \frac{\sum_{i=1}^{n} (S_i - O_i)}{n} \tag{21}$$

Maximum bias error (E_{max}) is the maximum difference between the observed and simulated cumulative values divided by position (n) on dataset (Eq. 22). By using dataset cumulative information, this error is able to indentify seasonal biases in the model – not only the general bias as in the case of MBE. When the model consistently overestimates or underestimates data, MBE and E_{max} should provide identical results. Otherwise E_{max} is a more robust metric than MBE and is useful to verify long-term errors, although it has less power to calculate daily or hourly errors.

$$E_{\max} = \left[\max_{i=1}^{n} \frac{\left(S_i^C - O_i^C\right)}{n^C} \right] \tag{22}$$

where S_i^C is the simulated cumulative values, O_i^C is the observed cumulative values, and n^C is extent of the accumulated dataset.

We calculate E_{max} (Eq. 22) for individual optimization at two sites. Additionally, we calculate E_{max} for DW multi-site OM by joining FNS and K77 data

that has been transformed in a single dataset. Finally, we compute SW multi-site as the average between the individual statistics.

Fig. 1.1, illustrates the subtle difference between E_{max} and MBE. Where E_{max} has the ability to find short-term error trends throughout the dataset, MBE calculates only the final error between simulated and observed data.

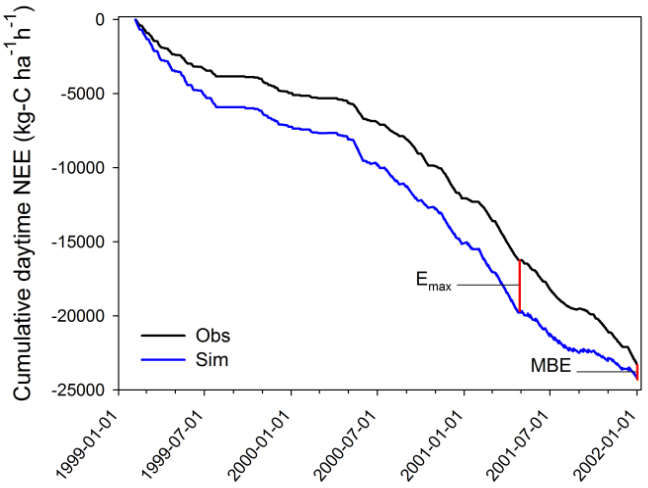

Fig. 1.1. Temporal differences between E_{max} (a) and MBE (b) adjustment measures.

1.3. RESULTS AND DISCUSSION

1.3.1. Correlation Coefficient

The correlation coefficient values range from 0.624 (FNS individual optimization) to 0.663 (K77 individual optimization) (Table 1.2). As expected, multi-site optimized values of *r* are in between the individually optimized values.

Table 1.2. Correlation coefficients (*r*) and parameters calibrated for FNS and K77 sites.

Site	OM	*r*	γ	*m*	V_{max} (µmol-CO_2 m^{-2} s^{-1})	St_m	K_u	K_f
FNS	Individual	0.624	0.010	10	120	-4	1.0	1.0
K77	Individual	0.663	0.010	10	120	-7	5.0	5.0
Multi-Site	DW	0.645	0.010	10	120	-4	1.0	1.0
Multi-Site	SW	0.642	0.010	10	120	-7	1.0	1.0

The parameters V_{max}, γ and *m* have the same values across all the optimization methods. However, St_m, K_u and K_f differ. γ and *m* depend only on the

plant species and V_{max} depends on the species and soil fertility. By contrast, St_m depends on soil physical characteristics and on rooting depth, which may in turn depend on soil depth or even how soil fertility varies with depth – all factors that can show considerable inter-site variability. K_u and K_f also depend on soil fertility and on local history of land use (fires, etc).

Our analysis indicates that when the correlation coefficient is optimized, the same values are obtained for the parameters that are species dependent. This suggests that parameters optimized using a single site may be representative for a wider region – at least for sites covered by the same plant species.

Parameters that depend on soil characteristics give different values for each individual optimization. Interestingly, the multi-site optimized parameters converge to either one of the values obtained from the individual optimizations.

Regression coefficients are very similar across optimization methods for FNS, with the slope ranging from 0.54 to 0.57 (Table 1.3, Fig. 1.2). For K77, there is much higher variation, with the gradient varying from 0.41 (multi-site DW) to 0.54 for the others OM. This indicates that when one of series is much shorter than the other ($n_{FNS} = 2.3\ n_{K77}$) the DW OM performs more poorly than the model for the site with the shorter time series. In contrast, the SW OM performs equally well for both sites.

Table 1.3. Linear regression of simulated vs. observed values of carbon flux for FNS and K77 sites.

Site	**OM**	***intercept***	***slope***	r^2
FNS	Individual	-2.544	0.5399	0.390
Multi-Site	DW	-2.544	0.5399	0.390
Multi-Site	SW	-2.642	0.5743	0.386
K77	Individual	-2.406	0.5398	0.439
Multi-Site	DW	-2.231	0.4125	0.409
Multi-Site	SW	-2.464	0.5390	0.439

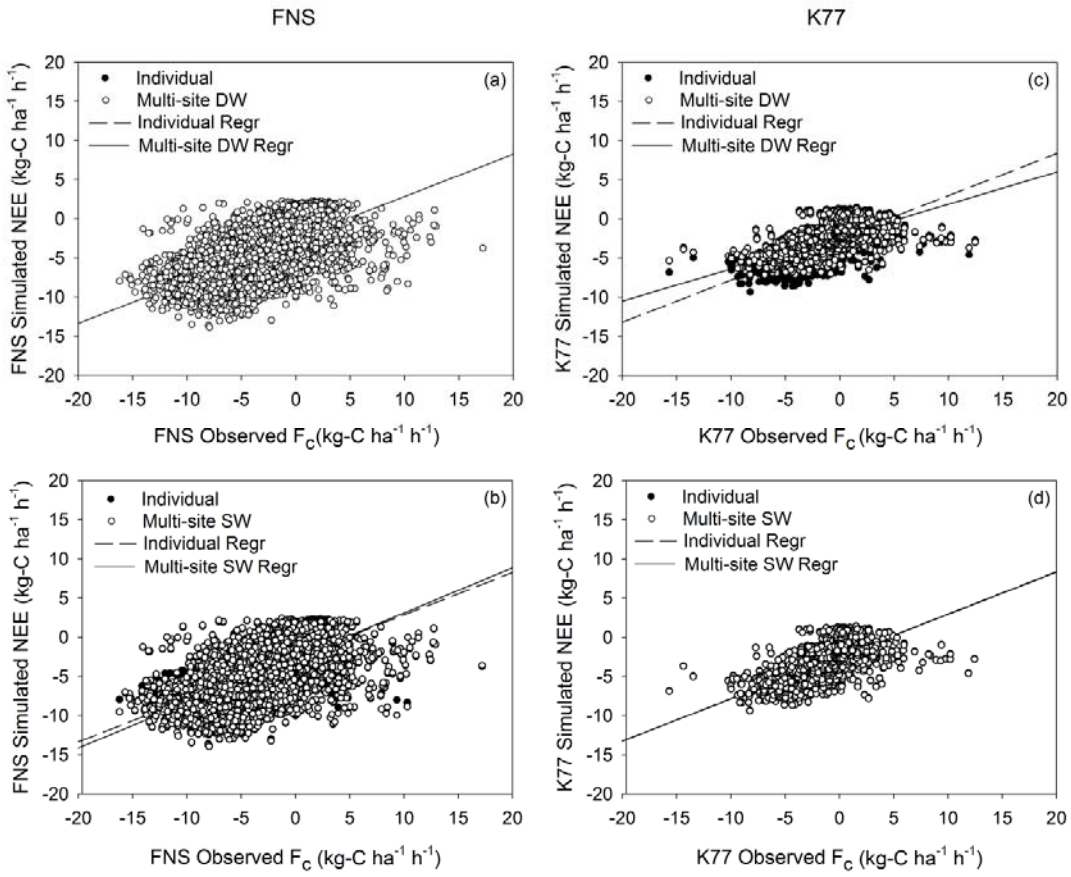

Fig. 1.2. Observed carbon flux and simulated NEE data to individual and multi-site optimization methods with data weight (DW) and site weight (SW) at the FNS site (a and b) and the K77 site (c and d).

1.3.2. Root Mean Squared Error

Optimization of RMSE produce values ranging from 2.16 kg-C ha^{-1} h^{-1} (K77 individual optimization) to 3.2 kg-C ha^{-1} h^{-1} (FNS individual optimization) (Table 1.4). Again, multi-site optimized values of RMSE fell between the individually optimized values.

Table 1.4. Root mean squared error (RMSE) and parameters calibrated for FNS and K77 sites.

Site	OM	*RMSE* (kg-C ha^{-1} h^{-1})	γ	m	V_{max} (μmol-CO_2 m^{-2} s^{-1})	St_m	K_u	K_f
FNS	Individual	3.20	0.010	10	95	-2	2.5	5.0
K77	Individual	2.16	0.010	10	120	-2	5.0	5.0
Multi-Site	DW	2.93	0.010	10	100	-2	5.0	5.0
Multi-Site	SW	2.69	0.010	10	100	-2	5.0	5.0

Estimated parameters display very little variation among the optimization methods – V_{max} being the most variable parameter. Both multi-site OM yielded the same value for V_{max} (100 μmol-CO_2 m^{-2} s^{-1}), which was predictably between the individually obtained values. This may indicate variation in soil fertility across sites.

Both SW and DW OM indicate the same set of parameters for optimum RMSE performance, although with different RMSE values. Regression coefficients are very similar in all sites, with the slope varying between 0.247 and 0.43 (Table 1.5).

Table 1.5. Linear regression of simulated vs. observed values of carbon flux to FNS and K77 sites.

SITE	OM	*intercept*	*slope*	r^2
FNS	Individual	-2.347	0.412	0.356
	DW	-2.341	0.430	0.362
	SW	-2.341	0.430	0.362
K77	Individual	-1.932	0.293	0.362
	DW	-1.840	0.247	0.334
	SW	-1.840	0.247	0.334

1.3.3. Mean Absolute Error

Optimized MAE values range from 1.57 kg-C ha^{-1} h^{-1} (K77 individual optimization) to 2.34 kg-C ha^{-1} h^{-1} (FNS individual optimization) (Table 1.6), with

multi-site optimized values between these two points. Again, γ and *m* are the same for all OM – a common behavior observed in all three dispersion minimization methods (*r*, RMSE and MAE). Moreover, in all three adjustment measures the optimized value is the same: 0.010 for γ and 10 for *m*. This strongly supports the argument that species specific parameters can be reliably optimized by any of these dispersion optimization methods. V_{max} varied between 100 and 115 μmol-CO_2 m^{-2} s^{-1} for each individual optimization, with multi-site optimization within this range. St_m also yielded the same values for all OM. As anticipated, the regression coefficients are similar to RMSE optimization (Table 1.7).

Table 1.6. Mean absolute error (MAE) and parameters calibrated for FNS and K77 sites.

Site	**OM**	***MAE*** **(kg-C ha^{-1} h^{-1})**	***γ***	***m***	V_{max} **(μmol-CO_2 m^{-2} s^{-1})**	St_m	K_u	K_f
FNS	Individual	2.34	0.010	10	100	-2	1.0	1.0
K77	Individual	1.57	0.010	10	115	-2	5.0	5.0
Multi-Site	DW	2.10	0.010	10	105	-2	5.0	5.0
Multi-Site	SW	1.96	0.010	10	105	-2	5.0	5.0

Table 1.7. Linear regression of simulated vs. observed values of carbon flux for FNS and K77 sites.

Site	**OM**	***intercept***	***slope***	r^2
FNS	Individual	-2.382	0.430	0.362
	DW	-2.348	0.447	0.367
	SW	-2.348	0.447	0.367
K77	Individual	-1.911	0.282	0.356
	DW	-1.865	0.259	0.342
	SW	-1.865	0.259	0.342

1.3.4. Mean Bias Error

The optimized MBE produces very small values (in the range of 10^{-4} kg-C ha^{-1} h^{-1}) for both the individual and multi-site DW OMs. In the case of the

multi-site SW OM, MBE is much higher, about 0.13 kg-C ha^{-1} h^{-1} (Table 1.8). The parameters selected by this objective function, which optimizes long-term adjustment of the model, all show high variability and are usually different from the ones obtained by dispersion optimization methods. Contrary to the previous model performance, when analyzed by the linear regression the model performs better at the FNS site (slopes in the range of 0.33 to 0.39) than at the K77 site (slopes in the range of 0.11 to 0.19) (Table 1.9). Although this adjustment measure provides a good adjustment for long-term performance, the hourly performance is poor.

Table 1.8. Mean bias error (MBE) and parameters calibrated for FNS and K77 sites.

Site	**OM**	***MBE*** (**kg-C ha^{-1} h^{-1}**)	***γ***	***m***	V_{max} (**µmol-CO_2 m^{-2} s^{-1}**)	St_m	K_u	K_f
FNS	Individual	0.0001009	0.040	9	65	-5	2.5	1.0
K77	Individual	0.0001011	0.040	8	35	-4	4.5	3.0
Multi-Site	DW	0.0001006	0.050	4	85	-1	5.0	1.5
Multi-Site	SW	0.1321681	0.030	10	90	-1	5.0	5.0

Table 1.9. Linear regression of simulated vs. observed values of carbon flux for FNS and K77 sites.

Site	**OM**	***intercept***	***slope***	r^2
FNS	Individual	-2.487	0.330	0.293
	DW	-2.282	0.341	0.273
	SW	-2.252	0.395	0.336
K77	Individual	-1.567	0.113	0.189
	DW	-1.823	0.179	0.198
	SW	-1.697	0.187	0.277

MBE is characterized by high variability in all estimated parameters and generates very different parameter values from the other adjustment measures. Furthermore, MBE obtained by the DW and SW multi-site optimizations are not within the range of the parameters obtained by individual calibration. Therefore, optimizing MBE does not seem to yield consistent model parameters.

1.3.5. Maximum Bias Error

Optimized E_{max} values range from 0.062 (DW multi-site OM) to 0.173 (K77 individual OM) (Table 1.10). Multi-site optimized values of E_{max} are once again between the individually optimized values. Estimated parameters exhibited low variability among the optimization methods. A test to ascertain if the order in which the data were accumulated (FNS-K77 or K77-FNS) on the DW multi-site OM influenced the results demonstrated that it made no difference if the accumulation started with K77 or FNS.

Table 1.10. Maximum bias error (E_{max}) and parameters calibrated for FNS and K77 sites.

Site	OM	E_{max} (kg-C ha^{-1} h^{-1})	γ	m	V_{max} (µmol-CO_2 m^{-2} s^{-1})	St_m	K_u	K_f
FNS	Individual	0.090	0.015	4	115	-1	5.0	5.0
K77	Individual	0.173	0.010	10	85	-1	1.0	2.0
Multi-Site	DW FNS-K77	0.062	0.015	7	95	-1	1.0	4.5
Multi-Site	SW	0.159	0.010	10	90	-1	5.0	3.5

The linear regression performs better at the FNS site (slopes in the range of 0.37 - 0.38) than at K77 (slopes in the range of 0.17 - 0.18). Determination coefficients range from 0.28 to 0.33 (Table 1.11), FNS having larger values than K77. Among the multi-site optimization methods there are no significant differences in the r^2 values.

Table 1.11. Linear regression of simulated vs. observed values of carbon flux for FNS and K77 sites.

Site	OM	*intercept*	*slope*	r^2
FNS	Individual	-2.280	0.381	0.311
	DW FNS-K77	-2.273	0.378	0.332
	SW	-2.210	0.370	0.335
K77	Individual	-1.641	0.169	0.281
	DW FNS-K77	-1.730	0.183	0.276
	SW	-1.640	0.179	0.289

1.3.6. Assessment of Multi-Site Optimization

Initially, the SW OM and DW OM are compared across all adjustment measures. The DW multi-site OM has higher RMSE and MAE than the SW multi-site OM (Table 1.12). This result is caused by the higher squared errors at FNS (70% of the data) than K77 during individual optimization. Hence the weighted average is higher than the sites mean. In contrast, FNS has lower MBE and E_{max} values than K77 when individually optimized. For the same reason of differences in data length, the DW multi-site OM is lower than SW multi-site OM. Finally, correlation coefficients are similar in both SW and DW OMs.

Table 1.12. Multi-site optimization methods and adjustment measures values.

Optimization	***r***	***RMSE*** **(kg-C ha^{-1} h^{-1})**	***MAE*** **(kg-C ha^{-1} h^{-1})**	***MBE*** **(kg-C ha^{-1} h^{-1})**	E_{max} **(kg-C ha^{-1} h^{-1})**
Multi-Site DW	0.645	2.93	2.10	0.0001006	0.062
Multi-Site SW	0.642	2.69	1.96	0.1321681	0.159

A comparison of DW and SW multi-site OMs reveals an apparent weakness in the DW multi-site OM – if the site with more data has a poor individual performance in the calibration, then the performance of the optimized DW multi-site model necessarily will also be poor. The SW multi-site OM gives the same weight to both sites, so the errors are evenly distributed. Based on these results we would argue that the SW multi-site OM is generally a better optimization method than DW multi-site OM.

Another notable feature of this study is that adjustment measures that optimize long-term model performance – MBE and E_{max} – produce high inter-parameter variability. Adjustment measures that optimize model dispersion (RMSE, MAE and *r*) may or may not produce parameters with the same value or with small

variability. The physiological parameters that depend mainly on the plant species (γ, m, V_{max} and St_m) either show constant values across OM or small variability. However, parameters (K_u and K_f) that depend mainly on soil conditions (soil bacteria, soil fertility, fire history) show higher variability across sites and OMs.

It is also important to discuss the range of parameter values obtained by various objective functions (Table 1.13): Optimized values of V_{max} vary widely from low values (35 μmol-CO_2 m^{-2} s^{-1}, K77 individual calibration using MBE as objective function), to high values (120 μmol-CO_2 m^{-2} s^{-1}, using correlation coefficient as adjustment measure). All optimization methods using r produce identical values for V_{max} – 120 μmol-CO_2 m^{-2} s^{-1}. The use of RMSE and MAE for multi-site OMs produce parameter values within the range of the individual optimization. In contrast, MBE and E_{max} generate very distinct values. In the case of E_{max}, multi-site optimized parameters are in the range of the single-site optimized parameters. MBE produces inconsistent V_{max} parameters, in the sense that multi-site optimized parameters are not in the range of the single-site optimized parameters. These results are very different from the values reported by Kucharik et al. (2000): V_{max} = 15 μmol-CO_2 m^{-2} s^{-1} for the C4 grass for the IBIS model. This difference is likely to arise from the variation among sites used in both analysis, as there is considerable variation in plant communities and soil.

Optimized values of m vary from 4 to 10. It should be noted that r, RMSE and MAE generate identical values for m (10) and the multi-site OM produces parameters values in the range of the single-site optimized parameters. In contrast, MBE retrieves different values (ranging from 4 to 10) and the multi-site optimized parameters are not in the range of the single-site optimized parameters – an inconsistent result. E_{max} values also range from 4 to 10. The multi-site OM generates

parameter values in the range of the individual optimization. Once again, our results differ from those of Kucharik et al. (2000), who found m = 4 for the C4 grass for the IBIS model. This is likely to be an individual response to sites, and shows the importance of site-specific calibration of the sensitive individual parameters rather than using the default on IBIS.

Table 1.13. Variability of parameters calibrated by all objective functions.

Parameters Calibrated	**OM**	***r***	***RMSE*** **(kg-C ha^{-1} h^{-1})**	***MAE*** **(kg-C ha^{-1} h^{-1})**	***MBE*** **(kg-C ha^{-1} h^{-1})**	E_{max} **(kg-C ha^{-1} h^{-1})**
V_{max} µmol-CO_2 m^{-2} s^{-1}	FNS	120	95	100	65	115
	K77	120	120	115	35	85
	DW	120	100	105	85	95
	SW	120	100	105	90	90
m	FNS	10	10	10	9	4
	K77	10	10	10	8	10
	DW	10	10	10	4	7
	SW	10	10	10	10	10
γ	FNS	0.010	0.010	0.010	0.040	0.015
	K77	0.010	0.010	0.010	0.040	0.010
	DW	0.010	0.010	0.010	0.050	0.015
	SW	0.010	0.010	0.010	0.030	0.010
St_m	FNS	-4	-2	-2	-5	-1
	K77	-7	-2	-2	-4	-1
	DW	-4	-2	-2	-1	-1
	SW	-7	-2	-2	-1	-1
K_u	FNS	1.0	2.5	1.0	2.5	5.0
	K77	5.0	5.0	5.0	4.5	1.0
	DW	1.0	5.0	5.0	5.0	1.0
	SW	1.0	5.0	5.0	5.0	5.0
K_f	FNS	1.0	5.0	1.0	1.0	5.0
	K77	5.0	5.0	5.0	3.0	2.0
	DW	1.0	5.0	5.0	1.5	4.5
	SW	1.0	5.0	5.0	5.0	3.5

RMSE, r and MAE generate the same value of γ (0.010), with the multi-site optimized parameters in the same range as the optimized single-site parameters. In the case of MBE, γ range from 0.030 to 0.050 with the optimized multi-site parameters outside the range of the optimized single-site parameters. E_{max} has values ranging from 0.010 to 0.015, with the multi-site OM parameters values in the range of the individual optimization.

Optimized values of St_m vary from -1 to -7. Using r, multi-site OM find parameter values within the range of the individual optimization. RMSE and MAE generate the same parameters values (-2) as individual and multi-site methods. MBE vary from -1 to -5, and the optimized multi-site parameters are outside the range of the optimized single-site parameters. E_{max} generates the same parameters values (-1) for individual and multi-site methods.

Optimized values of K_u range from 1.0 to 5.0. r has the same variation, with the multi-site OM finds parameters in the range of the individual optimization. RMSE and MAE find very similar values and the multi-site OM finds parameters values in the range of the optimized single-site parameters. MBE to the optimized multi-site parameters are not in the range of the optimized single-site parameters, demonstrating inconsistent results. In the case of E_{max}, optimized multi-site parameters are in the range of optimized single-site parameters.

Optimized values of K_f range from 1.0 to 5.0 with RMSE and MAE generating similar values. All adjustment measures for the multi-site OM, with the exception of MBE, generate parameter values within the range of the optimized single-site parameters. E_{max} values were in the range of 2.0 to 5.0.

Recent studies (e.g. Alton, 2011; Groenendijk et al., 2011) have questioned the validity of model parameters within the PFT classification. Groenendijk et al.

(2011) observe that a PFT classification introduces uncertainty to the short term variation of photosynthesis and transpiration fluxes, and therefore argue that model parameters are more variable than assumed with PFTs. Alton (2011) reports that carbon fluxes are somewhat sensitive to the assumed parameter value distribution within each PFT and observed that plant parameters retrieved from model optimization show some disaggregation into PFTs, but that the overlap between PFTs is generally large – as is found in field measurements.

Although we agree that parameters vary within a PFT, we designed this optimization experiment to try to eliminate (or at least minimize) this variability through working with sites covered by the same plant species. Nevertheless, the choice of objective function still introduces uncertainties into the optimized set of parameters. The results of this study support the notion that different objective functions generate different parameters solutions, and that it is therefore difficult to find a single global optimum parameter solution (Zhang, 2008). Depending on the adjustment measure or objective function used, different parameter solutions are found because each one emphasizes distinct and different statistical aspects: RMSE, MAE and *r* minimize the dispersion of the data while MBE and E_{max} maximize long-term adjustment of the models.

Willmott and Matsuura (2005) report that MAE is the most natural measure of average error, and should therefore be the basis for evaluations of model performance error. In the present study we do not find significant differences in the performance of RMSE and MAE. In turn, *r* is not always a good adjustment measure. The data may have a small dispersion but, as the slope coefficient of the regression between simulated and observed data is usually not close to 1, it is possible to have a good correlation coefficient with a poor simulation.

MBE also presents some problems. First, it does not detect the difference that occurs between data averages for the intermediate period of the dataset (Fig. 1.1). Second, parameters optimized by multi-site calibration are not representative, because they are not within the range of the individual calibration values. E_{max} is a more robust adjustment measure to evaluate long-term performance, because it can detect the bias in intermediate periods as well as at the end of the dataset. The best performance of the model with the parameter set obtained by E_{max} optimization is clearly shown in Figures 1.3 and 1.4. However, the results vary considerably with site. At FNS the model performs well throughout the simulation period for both optimized MBE and optimized E_{max} runs (Fig. 1.3a and 1.4a). By contrast, the results are remarkably different at K77 where the observed data change slope at the end of the simulation period – which the model cannot reproduce correctly in any case (Fig. 1.3b and 1.4b). The individual MBE-optimization picks a set of parameters that make the simulated cumulative curve meet the observed one exactly at the end of the simulation period (Fig. 1.3b), while the E_{max} optimization generates a set of parameters that results in a more consistent performance throughout the simulation. Moreover, E_{max} always obtained consistent values for the calibrated parameters, i.e., the optimized multi-site parameters are always within the range of the individual optimization parameters.

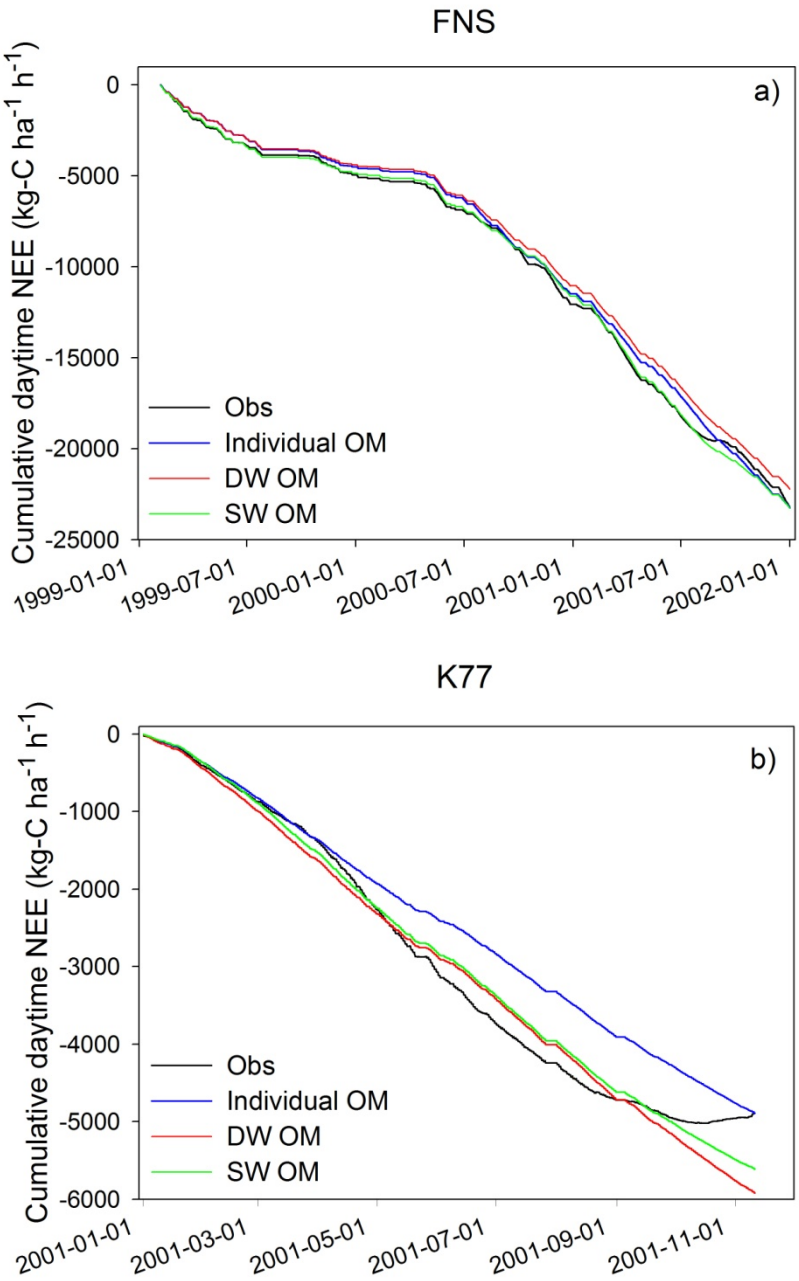

Fig. 1.3. NEE accumulated for MBE, to observed data and to each optimization method.

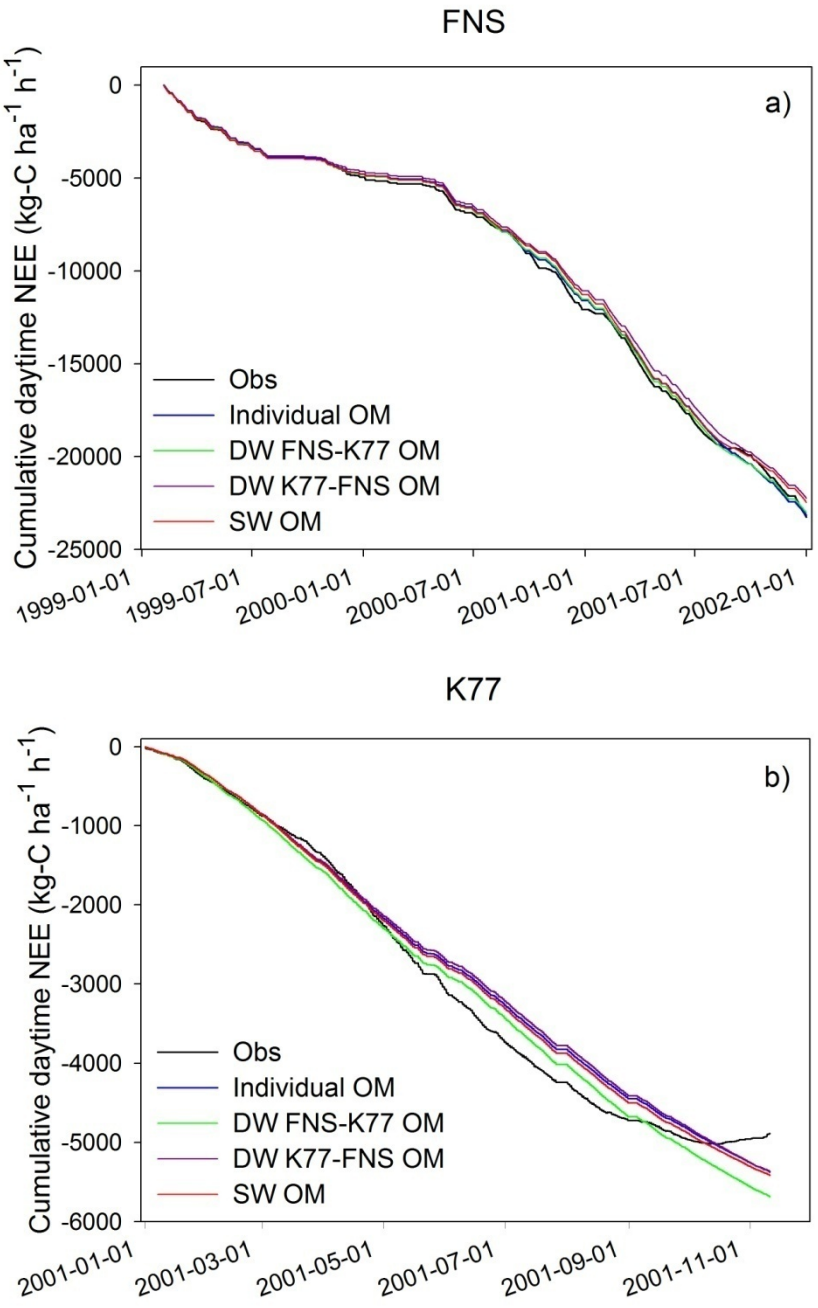

Fig. 1.4. NEE accumulated for E_{max}, to observed data and to each optimization method.

1.4. CONCLUSIONS

The recent availability of flux data for sites that are covered by the same PFT opens new research possibilities for the optimization of land surface models, allowing the test of commonly used assumption that parameters obtained at a single site could be representative of all sites covered by that PFT.

Here, we use two sites covered by the same species to eliminate a major source of parameters variability across sites – ecosystem composition. However, other sources of parameter variability, such as soil fertility, soil depth, method of clearing and fire history, could not be controlled.

The results produced by 20 objective functions (five adjustment measures subject to four optimization methods for two sites) are evaluated to investigate the consistency of the results in a multi-site model optimization. The results suggest that the (frequently used) choice of MBE as adjustment measure does not produce consistent parameters across sites and should be discarded in favor of one of the other adjustment measures that generally generate consistent parameters. We also conclude that assigning the same weight for each site (SW) is a better practice than weighting model performance according to the length of each data series (DW).

Overall, our recommendation for the choice of objective function depends on the intended use of the model. If the model is intended for short term studies (for

example, on weather forecast applications) where the accurate simulation of hourly fluxes is more important and dispersion needs to be minimized, then we recommend using the site-weighted mean absolute error (SW MAE). This objective function produced the most consistent results, specifically: V_{max} = 105 μmol-CO_2 m^{-2} s^{-1}, m = 10, γ = 0.010, St_m = -2, K_u = 5.0 and K_f = 5.0.

Conversely, if the model is going to be used on longer time scales (such as IPCC simulations) where long-term adjustment is more important, then we recommend using the site-weighted maximum bias error (SW E_{max}) as an objective function. In this case, the model optimized parameters are V_{max} = 90 μmol-CO_2 m^{-2} s^{-1}, m = 10, γ = 0.010, St_m = -1, K_u = 5.0 and K_f = 3.5.

This research provides insights into some of the main questions as it improves our understanding about multi-site variation in the specific parameters used in ecosystem modeling. Parameters like γ and m can be optimized using data from a single site, however, variables such as V_{max} and others frequently need to be optimized from multiple sites because they vary between plant species and according to soil conditions.

The results of this study indicate that it is possible to perform multi-site calibration in land surface models and still obtain consistent results for both sites. Although multi-site calibration may become more common in future, the complex problem of application to heterogeneous sites remains. When ecosystem composition is also a source of parameter variability, we may need to obtain two model parameters statistics: parameter mean and parameter standard deviation. These two statistics would allow a stochastic representation of PFTs, potentially leading to the use of Monte Carlo simulations of ecosystems at the regional scale. Obtaining these

statistics would require at least three eddy flux tower sites for the same ecosystem and the development of new methodologies.

CHAPTER 2

ESTIMATION OF NET PRIMARY PRODUCTION IN GRASSLAND AND CERRADO IN SOUTH AMERICA

2.1. INTRODUCTION

Since the industrial revolution the global average concentration of CO_2 has risen from 285 ppm to 379 ppm in 2005 (Forster et al., 2007) and the current concentration is the highest in the last 650,000 years (Petit et al., 1999; Siegenthaler et al., 2005). This large increase in CO_2 is mainly due to burning fossil fuel and significant land use changes across the globe, the most important form of which is deforestation that is frequently associated with biomass burning (Denman et al., 2007).

The impacts of an increased concentration of atmospheric CO_2 on regional climate and terrestrial ecosystems are complex and bi-directional. For example, increased CO_2 can cause significant variations in the climate that in turn influence terrestrial carbon sequestration (Nemani et al., 2003). Thus, global carbon balance is influenced by both changes in the rates of anthropogenic emissions and dynamics of many terrestrial and oceanic processes (Sabine et al., 2004; Canadell et al., 2007); Such processes can result in the emission or removal of CO_2 from the atmosphere. For example, 24% of total anthropogenic emissions of carbon into the atmosphere (during the period 2000 to 2006) were absorbed by the oceans, and 30% were absorbed by terrestrial ecosystems (Schimel et al., 2001; Canadell et al., 2007).

However, the specific geographic locations of this carbon absorption are not clearly known (Friend et al., 2007; Zaks et al., 2007).

To better understand variation in the terrestrial carbon balance it is necessary to perform fundamental research on global carbon cycle dynamics. In this respect, primary production has received considerable attention in recent years because its role in the global carbon cycle is clearly understood (Zhao et al., 2005) and it is one of the most important contributing variables for terrestrial carbon balance. Carbon sequestration rate in an ecosystem, or NPP (net primary production), is defined as the net flux of carbon to plants through photosynthesis. It is equal to the difference between gross primary production (GPP) and autotrophic respiration (R_a), integrated over time. NPP rate, together with mortality rate, are also two of the main determinants of ecosystem structure. Moreover, there are several factors that are known to influence NPP, such as climate, topography, soil, plant, microbial characteristics, disturbance and anthropogenic impacts (Field et al., 1995).

Field estimates of NPP in which biomass increments (leaves, stems and roots) are monitored over time is an expensive and labor intensive methodology and is typically performed in experimental sites of relatively small dimensions. Despite the effort involved in obtaining these measures, they are insufficient for scaling up for regional scale estimates or for monitoring over time. However, there are other ways to estimate NPP – most notably through biosphere-atmosphere interaction models or by remote sensing techniques. These methods are the most frequently used to quantify spatio-temporal variability in NPP (Field et al., 1995; Running et al., 1999; Running et al., 2004; Zhao et al., 2005; Pinker et al., 2010; Ise et al., 2010; Wang et al., 2011).

Remote sensing data can provide valuable and continuous space-time information of several biophysical processes, disturbance and terrestrial biosphere dynamics, as well as information about the impacts of climate change on terrestrial ecosystems. The MODIS (Moderate Resolution Imaging Spectroradiometer) - initially launched in 1999 aboard the Terra satellite with a second sensor launched in 2002 aboard the Aqua satellite - initiated a new era in globally monitoring climate change. MODIS is considered the most important tool for remote sensing of the global environmental, due to its complete coverage of the Earth's surface. Various products are generated by MODIS including the vegetation index (NDVI, Normalized Difference Vegetation Index), leaf area index (LAI), fraction of absorbed photosynthetic active radiation (FAPAR), and land surface changes. The most relevant product for studies of carbon cycling is the NPP product, which is available every 8 days at a 1 km^2 spatial resolution.

Many of the MODIS products are still in the process of validation (e.g. Running et al., 2004; Senna et al., 2005; Cohen et al., 2006; Plummer, 2006; Turner et al., 2006; Fensholt et al., 2006; Yuan et al., 2011). The validated product can be used for monitoring but, given the limitations of the validation, cannot be used to make effective and accurate future predictions. In this context, the use of biosphere-atmosphere interaction models which capture broad-scale ecosystem processes has much greater applicability for forecasting future trends in key variables such as NPP. For this reason several biosphere-atmosphere interaction models have been extensively used in monitoring and forecasting NPP or NEE (net ecosystem exchange), such as the Miami Model (Lieth, 1973), CASA (Potter et al., 1993), IBIS (Foley et al., 1996), RATE (Nunes, 2008) and VPRM (Mahadevan et al., 2008).

Many of these models use information from remote sensing data to improve their estimates (Huang et al., 2010; Nayak et al., 2010; Wang et al., 2011).

There are also several ongoing studies using biosphere models to assess the impact of increasing CO_2 on carbon cycle in terrestrial ecosystems (Friend et al., 2007; Piao et al., 2008; Sitch et al., 2008; Luyssaert et al., 2010). Tropical rainforests have been the main focus of many of these studies because they are considered to have very high potential for carbon assimilation. This focus is understandable, but neglects the role of other important ecosystems in driving climate change. For example, the cerrado is the second largest Brazil biome and is suffering from rates of habitat loss far greater than the Amazon rainforest and has already lost more than 50% of its natural vegetation cover (Sano et al., 2008). Agricultural pasture is another important land use which has so far in studies of carbon cycling, despite that fact that it is the main land use category replacing rain forest and cerrado (Costa and Pires, 2010).

One of the main advantages of extending models to other major biomes is better regional NPP estimates which, in turn, contribute to knowledge of the global carbon cycle and carbon storage. In this article, we attempt to fill this gap in understanding by creating realistic simulations of carbon balance for areas of cerrado and grassland in South America. We do this by using the RATE Algorithm to estimate spatio-temporal patterns of NPP over grassland and cerrado using land cover, LAI and FAPAR datasets derived from the MODIS product. We then compare the simulated NPP from the biosphere-atmosphere interaction model with estimates from the MODIS-NPP algorithm (Running et al., 2004).

2.2. MATERIAL AND METHODS

2.2.1. RATE Algorithm: Description

RATE is an algorithm for monitoring and quantifying the NPP of terrestrial ecosystems of South America (full specification is provided in Nunes, 2008). Basically, RATE is derived from MODIS remote sensing data of km^2 resolution of land cover, leaf area index (LAI) and fraction of absorbed photosynthetic active radiation (FAPAR), and data reanalysis of meteorological variables provided by the NCEP/NCAR (National Centers for Environmental Prediction / National Center for Atmospheric Research). The RATE algorithm is currently able to provide estimates of temporal and spatial carbon fixation rate to tropical forest ecosystem, from 8 in 8 days, with a spatial resolution of 0.05 degrees.

2.2.1.1. SITE Model

An essential component of the RATE algorithm is the incorporation of a simple biosphere-atmosphere interaction model to calculate NPP. The algorithm uses the simple tropical ecosystem model (SITE) developed by Santos and Costa (2004) which is based on a number of earlier models, mainly LSX (Pollard and Thompson,

1995), LSM (Bonan, 1996), IBIS (Foley et al., 1996) and SiB2 (Sellers et al., 1996). SITE incorporates several biophysical and biochemical processes, such as canopy infrared radiation balance, solar radiation balance, aerodynamic processes, canopy physiology and transpiration, water balance intercepted by the canopy, mass and energy transport in the atmosphere, soil heat flux, soil water flux and carbon balance. Although the model is simplified, it has the complexity necessary to represent the dynamic of ecosystems such as the physiological processes involved in carbon cycling (Santos and Costa, 2004).

SITE is a dynamical point model that uses an integration time step (*dt*) of one hour and is structured with one canopy layer and two soil layers. The model also incorporates various biophysical characteristics of the vegetation such as specific leaf area and stem area index (see Santos and Costa, 2004, for full specification). The main output variables of the model are latent heat flux, sensible heat flux, water vapor flux and net exchange ecosystem and net primary production.

The carbon fixed by photosynthesis (NPP) can be allocated in four different reservoirs of the plants: leaves, stems, fine roots and coarse roots. The NPP (kg-C m^{-2} s^{-1}) is expressed as a function of the gross photosynthesis (A_g, mol-CO_2 m^{-2} s^{-1}) and canopy autotrophic respiration (leaves, stems and roots).

$$NPP = (1-\eta)\int (A_g - R_u - R_S - R_f - R_r)dt \quad (1)$$

where η = 0.3 is the fraction of carbon lost due to growth; A_g (Eq. 2) is the minimum between J_e (gross photosynthesis rate limited by light) and J_c (gross photosynthesis rate limited by the activity of the Rubisco enzyme); R_u is the respiration rate of leaves (mol-CO_2 m^{-2} s^{-1}), R_s is the respiration rate of stems, R_f is the respiration rate of the fine roots and R_r is the respiration rate of the thick roots.

$$A_g = \min(J_e, J_c) \quad (2)$$

J_e is expressed by:

$$J_e = \alpha_4 APAR\left(\frac{CO_{2i} - \Gamma}{CO_{2i} + 2\Gamma}\right) \quad (3)$$

where α_4 is the intrinsic quantum efficiency for C_4 plants (mol CO_2 mol^{-1} photons); *APAR* is the density flux of the absorbed photosynthetically active radiation (mol photons m^{-2} s^{-1}); CO_{2i} is the CO_2 concentration in the intercellular spaces of the leaves; and $\Gamma = [O_2]/2\tau$ is the compensation point for gross photosynthesis (mol mol^{-1}); O_2 is the oxygen concentration and τ is the CO_2/O_2 ratio of kinetic parameters.

J_c is given by:

$$J_c = \frac{V_m\left(CO_{2i} - \Gamma\right)}{CO_{2i} + K_c\left(1 + \left([O_2]/K_0\right)\right)} \quad (4)$$

where K_c and K_0 are two levels of the Michaelis constant: K_c is used to fix CO_2 and K_0 to inhibit oxygenation (mol mol^{-1}).

$$V_m = V_{\max}\, T_{vm} St \quad (5)$$

where V_m is the maximum Rubisco enzyme capacity (V_m, μmol-CO_2 m^{-2} s^{-1}) for the carboxylase function at 15ºC, calculated through the V_{max} parameter and modified by soil moisture stress (*St*) and temperature stress (T_{vm}) functions.

NEE is expressed as the difference between soil heterotrophic respiration (R_H) and the net primary production (NPP). Negative values of NEE indicate assimilation of carbon by the ecosystem and positive values indicate carbon release by the ecosystem.

$$NEE = R_H - NPP \quad (6)$$

In this research, the SITE model uses regional scale data that are assimilated into the model. The assimilated data are obtained from: the MODIS land cover

product (MOD12Q1, collection4), the MODIS LAI/FAPAR product (MOD15A2, collection 5); and meteorological variables from NCEP/NCAR re-analysis. A detailed description of data collection is given below.

2.2.1.2. Land Cover data

Land cover data were taken from MODIS products (MOD12Q1, collection 4), over 4 years, from 2001 to 2004. Thus, in order to generate a decade long data series, data from 2001 were replicated for 2000 and data from 2004 were replicated for 2005 to 2010. The land cover product uses 17 vegetation classes, with a resolution of 1 km^2. In this research we utilize the data for cerrado and grassland vegetation classes.

Three steps are required to transform data from the MOD12Q1 product into a form that can be input into the RATE Algorithm. First, we use the MODIS Reprojection Tool to agregate 31 tiles, and to create a single archive corresponding to South America. Initially this archive has a spatial resolution of 1 km^2 in Senoidal projection, and therefore was converted into geographical coordinates with a spatial resolution of 0.08^o for cerrado and 0.05^o for grassland.

Second, we use a simple script in NCL (NCAR Command Language) to convert files in HDF-EOS format from MOD12Q1 (obtained for the spatial resolution of 0.08^o for cerrado and 0.05^o for grassland) into the network Commom Data Form (netCDF) format. This format supports creation, access and sharing of vectorially stored scientific data.

Finally, we generated cerrado and grassland maps for South America with a spatial resolution of 0.08^o and 0.05^o, respectively. Thus, a single file in a netCDF format is created that contains data with pixels of biomes used, where each pixel is

quantified in terms of the frequency of occurrence of the required biome. The RATE Algorithm considers only pixels of land cover with the frequency of occurrence of the biome greater than or equal to 60% for cerrado and 40% for grassland.

2.2.1.3. LAI and FAPAR data

Two important variables used on calculation of NPP, LAI and FAPAR data were taken at Collection 5 data products from MODIS (MOD15A2). LAI and FAPAR products have spatial resolution of 1 km^2 and temporal resolution of eight days. These data were acquired for the period March 2000 to December 2010 and converted into a spatial resolution of 0.01^o in netCDF data format. The algorithm used to generate MODIS LAI and FAPAR product explores the information contained in the MODIS atmospherically corrected surface reflectance in seven spectral bands.

We create the 0.05^o and 0.08^o grid using the maximum values of LAI and FAPAR at 0.01^o using original data for grassland and cerrado. The RATE algorithm has factor corrections for LAI and FAPAR that are obtained from the observed error. The correction factor is equal one minus relative error average, where the relative error is calculated the difference between LAI observed and LAI obtained from MODIS. The correction factor is equal to 0.6338 to grassland and 0.6978 to cerrado. As we don't have FAPAR observed data for both cerrado and grassland sites, we use a correction factor that was estimated for tropical forest.

2.2.1.4. Meteorological Variables from Reanalysis

The SITE model is forced by a time-series of meteorological data. We used the re-analysis product of NCEP/NCAR (National Centers for Environmental Prediction / National Center for Atmospheric Research), with the following variables: incident short wave radiation, incident long wave radiation and precipitation, with 1.9º of spatial resolution, air temperature, relative humidity, horizontal components of wind speed (u and v) and atmospheric pressure, with 2.5º spatial resolution. These variables were acquired for the period March 2000 to December 2010, at an interval of 6 hours.

We convert these data through a bi-linear interpolation into a spatial resolution of 0.05º and 0.08º and to the required integration interval of 1 hour by linear interpolation.

2.2.2. Calibration

2.2.2.1. Experimental sites and field measurements

The model was calibrated at four experimental sites – two cerrado sites in Central Brazil, and two grassland sites located in the Amazon which are part of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) towers network.

The first cerrado site is located in the floodplain of Bananal Island (hereafter referred to as site BAN) (09°49'S, 50°09'W), in the state of Tocantins, Brazil. This site has very heterogeneous vegetation, including a mixture of cerradão (tall woodland savanna with 18 m tall trees and sparse shrubs), cerrado (dense scrub with

5 m tall trees and grass understory) and campo (natural grassland). The second cerrado site is located in southeastern Brazil, in the Pé-de-Gigante reserve (hereafter referred to as site PDG) (21°37'S, 47°39'W), in the state of São Paulo, Brazil. Its vegetation consists of cerrado *sensu strictu*, a closed scrub with 5 – 10 m tall trees and a dense herbaceous understory (da Rocha et al, 2009).

The first grassland site is located at the Fazenda Nossa Senhora ranch (hereafter referred to as site FNS) (10^{o}45'S, 62°21'W, 230 m), in Ouro Preto d'Oeste, Rondônia State, Brazil. This site is in the center of a deforested area with an approximate radius of 50 km – deforestation was caused by an intentional fire in 1977 to clear land for crop cultivation. Since the early 1980s the area has been uniformly covered by the grass *Brachiaria brizantha*. Our second grassland site is on a farm at km 77 along the Santarém-Cuiabá highway (hereafter referred to as site K77) (03^{o}01'S, 54°53'W), near Santarém, Pará State, Brazil. The forest was cleared in 1990 and the field was planted with the same grass species as the first site, *Brachiaria brizantha*. Our analysis of K77 site is restricted to the above time periods because the site was burnt in November 2001 and subsequently converted to rice cultivation.

The model uses hourly meteorological input variables, such as air temperature, specific humidity, wind speed, incident short wave radiation, incident long wave radiation and precipitation. In the two grassland sites these data were measured from January 1999 to December 2001 and from January to November 2001, in the FNS and K77 sites, respectively. For two cerrado sites the input data were measured from January 2001 to December 2003 (PDG) and from January 2004 to December 2006 (BAN). Hourly carbon flux data was used to optimize NEE output of the SITE model. Carbon flux data for grassland sites were collected from February

1999 to December 2001 (a total of 6,264 data points) at the FNS site, and from January to November 2001 (a total of 2,768 data points) at the K77 site. Carbon flux for cerrado sites were collected from February 2002 to December 2003 (a total of 5,620 data points) at the PDG site and from January 2004 to November 2006 (a total of 8,158 data points) at the BAN site. Data used in this study were provided by LBA (Large-Scale Biosphere-Atmosphere Experiment in Amazonia) and LBA-DMIP(LBA Model Intercomparison Project).

In order to avoid the excessive weight attributable to respiration parameters during the night, we use only data collected between 07.00 and 17.00 hours for NEE optimization. Moreover, we filtered the carbon flux data to eliminate clearly spurious values – for example values of carbon flux lower than -20 kg-C ha^{-1} h^{-1}.

2.2.2.2. Optimization method and adjustment measure

We follow the multi-site optimization method and adjustment measures described in Fischer et al. (Submitted) (Chapter 1). In summary, we use optimize over using site-weighted parameters and use maximum bias error (E_{max}) as the adjustment measure, because we performing long-term simulations. The same calibration performed to grassland sites was done to cerrado sites. We find the identical results of the grassland sites for cerrado sites, indicating the same conclusions. A full specification is given in Fischer et al. (Submitted) and is available on request. The parameters calibrated used are related on Table 1.

Table 2.1. Calibrated parameters used in Algorithm RATE.

Site	γ	m	V_{max} (µmol-CO_2 m^{-2} s^{-1})	St_m	Ku	Kf
Grassland	0.010	10	90	-1	5.0	3.5
Cerrado	0.0080	3	105	-7	1.0	1.0

2.2.3. Comparison of results

A precise validation of simulated NPP requires *in situ* measurements of NPP. Unfortunately we do not have these measurements for any site, so NPP estimates from the RATE Algorithm are compared with the estimates of the NPP MODIS product (MOD17A3, collection 5.1). NPP MODIS product (a NPP algorithm provide annual growth of the terrestrial vegetation) was acquired from the period of the 2001 to 2010.

MODIS NPP Algorithm (Running et al., 2004) estimates annual NPP is based in the following equations:

$$GPP = PAR \times FAPAR \times \varepsilon T_f VPD_f \quad (7)$$

where *GPP* is the gross primary production, ε is the light use efficiency factor which is originally parameterized modeled global terrestrial GPP. *PAR* is photosynthetic active radiation and *FAPAR* is the fraction absorbed of PAR. T_f are function of minimum air temperature and VDP_f are function of vapor pressure deficit.

$$PSN_{net} = GPP - R_{lr} \quad (8)$$

where PSN_{net} is the net photosynthesis in daily basis and R_{lr} is daily maintenance respiration of leaves and fine roots.

The NPP annual is calculated by:

$$NPP = \sum_{1}^{365} (PSN_{net}) - R_g - R_m \quad (9)$$

where R_g is the annual growth respiration and R_m is the maintenance respiration of live cells in woody tissues.

So, NPP is the annual sum daily of net photosynthesis minus the growth and maintenance respiration.

2.3. RESULTS AND DISCUSSION

2.3.1. Monthly NPP patterns

Monthly NPP patterns are analyzed for cerrado and grassland ecosystems. Monthly NPP estimates for cerrado for 2010 range from -0.05 to 0.15 kg-C m^{-2} $month^{-1}$ (Fig. 2.1). Two growing seasons can be identified: November to April, which exhibits high values of NPP, and May to October, which is characterized by low values of NPP, in wettest period and driest period respectively (Fig. 2.2). The lowest values of NPP are found in October, the final month of dry season, in the central and northwest regions. In contrast, the highest values of NPP are found during all months of the wet season, but especially in the middle of the wet season (January). A high magnitude of NPP occurs in the central and northeast of Brazil, and lower values in central-west and southeast regions. In February, NPP has the same distribution, but with lower values, as in January. March is more similar with January, with higher values than February. In April, NPP values are similar in central Brazil, with higher values in the northeast. In May, NPP begins to fall in the central region and over the following month, these lower values of NPP spread throughout central Brazil. The low values of NPP extend through the month of October, and reach the lowest values at the end of dry season, in the central, north

and northwest regions. In November and December, NPP begins to increase again throughout Brazil.

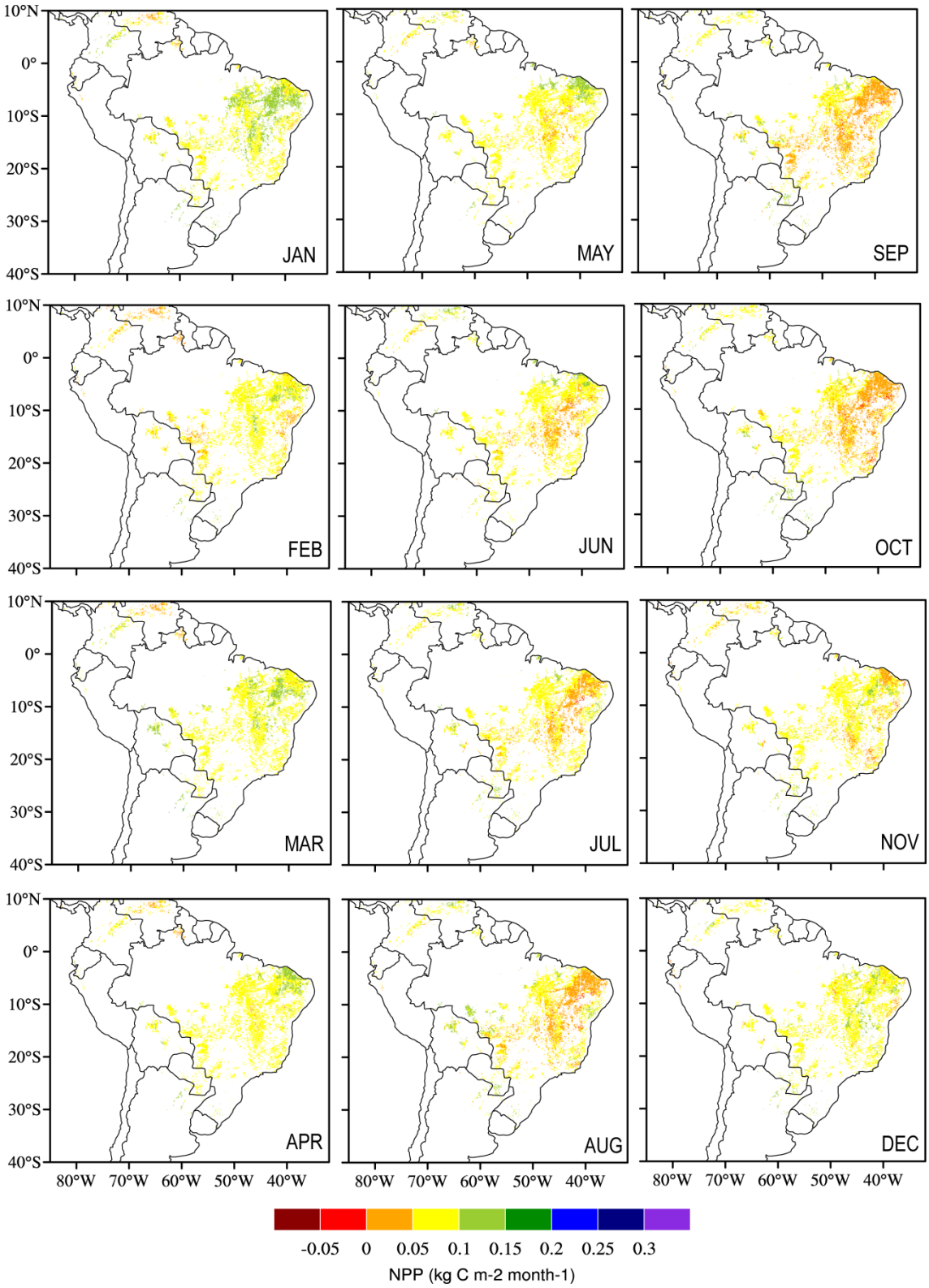

Figure 2.1. Cerrado spatial monthly NPP patterns (kg-C m^{-2} $month^{-1}$) in 2010.

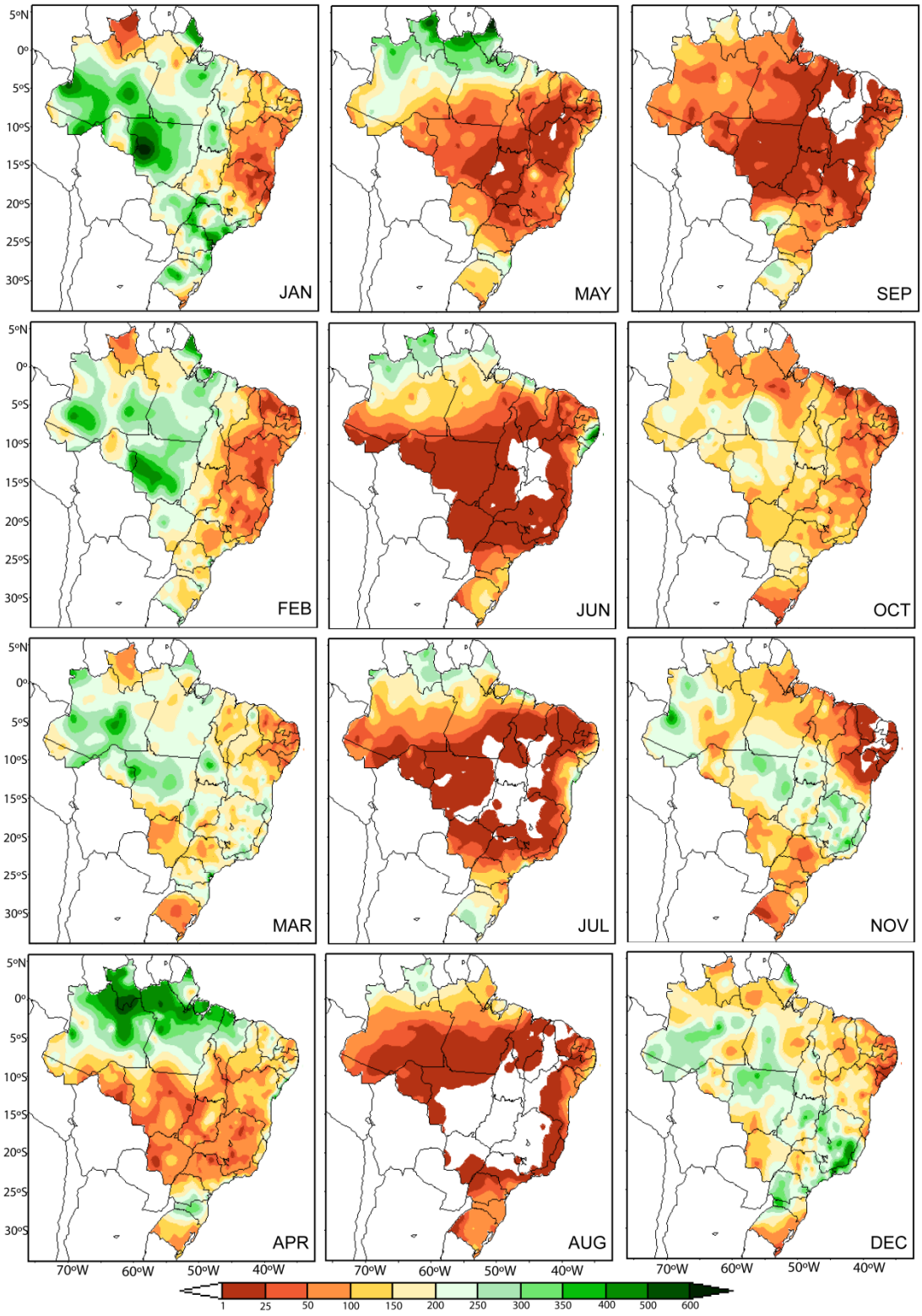

Figure 2.2. Accumulate monthly rainfall patterns (mm) in 2010.
Fonte: CPTEC.

NPP estimates for grassland in 2010 range from -0.05 to 0.15 kg-C m^{-2} $month^{-1}$ (Fig. 2.2). The grassland ecosystem is located in different parts of South America, including tropical and subtropical grassland, which have different behavior due to experiencing different climates. The tropical grassland has high values of NPP from November to April (wettest period) and has low values of NPP from May to October (driest period). Subtropical grassland is characterized by high values during the warm months of November to April and low values in the colder months (May to September). December has a different pattern, with lower values of NPP, probably due absence of rain since October until December in this region (Fig. 2.2). The west of the continent (Peru, Colombia and Venezuela) and in the central region of Argentina are characterized by low values of NPP during all months – except in the south of Peru which has high NPP values from January to April.

Analyzing every month individually, we can see that in January in central Brazil the values of NPP are very variable, showing both low and high values. In the south of Brazil, Uruguay, the northeast of Argentina and Paraguay, the magnitude of NPP is higher. In February NPP values are generally lower than in January. In March, the values of NPP increase again and then fall again in April. In May, with the coming of dry season in central Brazil and the coming of winter in the south of Brazil, Uruguay, Argentina and Paraguay, the magnitude of NPP decreases in all these regions. The reductions in NPP continue until October in central Brazil and until August in the south of Brazil and Uruguay. In Argentina the decrease in NPP continues until December. The lower values of NPP in November and December in the south Brazil and Uruguay are probably due to the low levels of precipitation that can occur in these months and which limit carbon fixation.

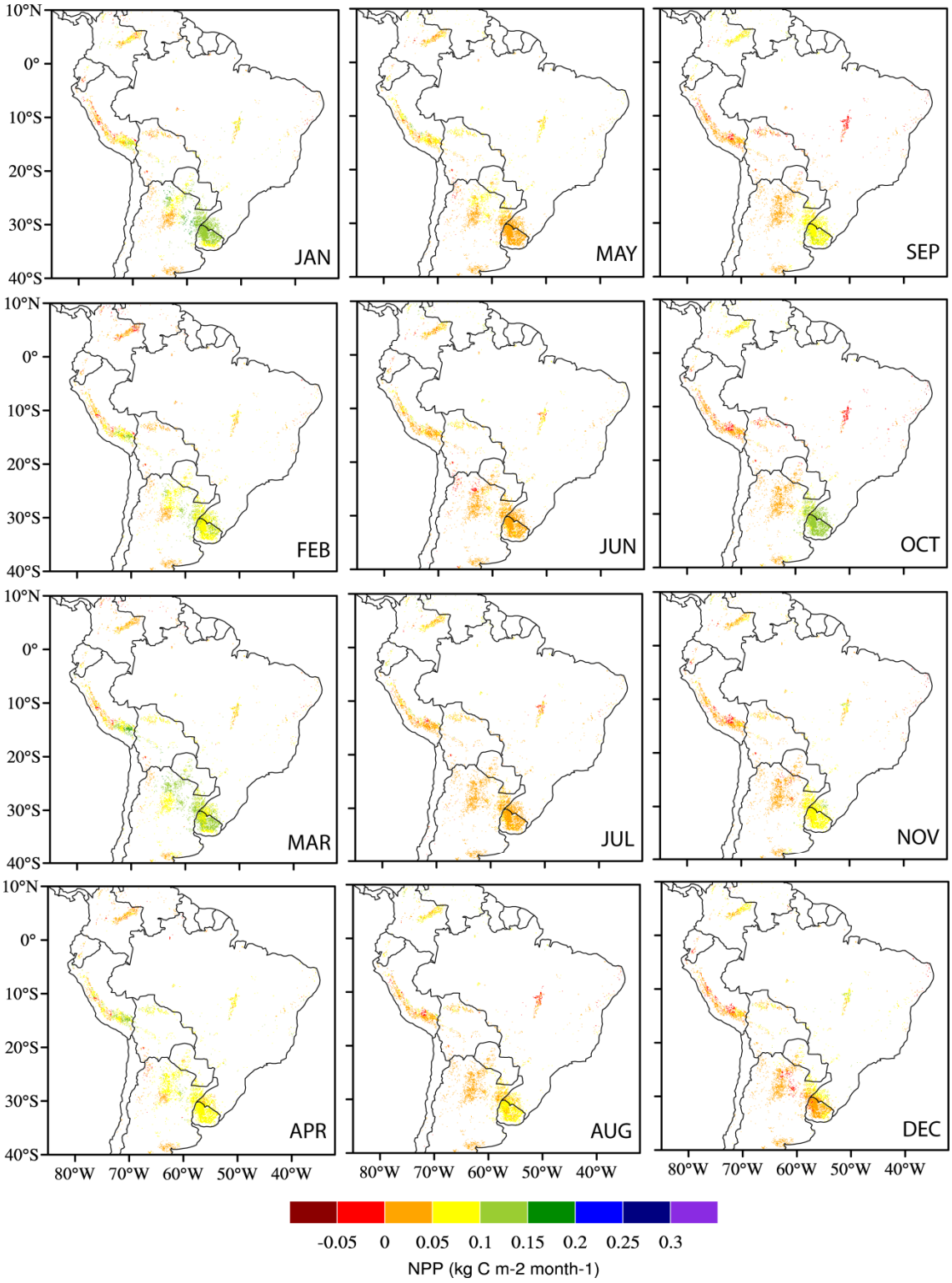

Figure 2.3. Grassland spatial monthly NPP patterns (kg-C m^{-2} $month^{-1}$) in 2010.

It is well known that carbon fixation is affected by climate and that plants respond to seasonality in the climate. In general, high values of NPP are found in places where neither water nor heat is a limiting factor. Conversely, low values of NPP are found in areas where limitations on the availability of some combination of water, heat and solar radiation reduce carbon fixation.

Cramer et al. (1999) affirm that carbon exchange is controlled by a combination of solar radiation and local environment conditions. Specifically, precipitation and temperature are the two major factors that govern the absorption of photosynthetically active radiation (PAR) and NPP of the biosphere. Luyssaert et al. (2007) report that temperature and precipitation account for 36% of the variability in observed NPP. The results presented here support this close linkage between NPP, temperature and precipitation, whereby NPP is inhibited by low temperature and precipitation.

Matsushita et al. (2004) argue that grassland NPP is particularly sensitive to changes in solar radiation, suggesting that this ecosystem is light-limited. However, Schloss et al. (1999) suggest that sensitivity of NPP estimates to solar radiation is dependent of latitude, which may explain the lack of strong patterns between solar radiation and NPP in grasslands in the current study.

In order to better assess the influence of seasonal variability we performed punctual estimations of NPP for PDG (cerrado) and FNS (grassland) experimental sites. The cerrado site showed an accentuated seasonal variation of NPP (Fig. 2.3), with high values from the end of spring until to end of summer / start of autumn (> 0.10 kg-C m^{-2} $month^{-1}$). The highest value of NPP is found in March (0.121 ±0.015 kg-C m^{-2} $month^{-1}$), and from this month the magnitude of NPP rapidly declines and reaches the lowest values in September (0.055 ± 0.028 kg-C m^{-2} $month^{-1}$). Values of

NPP clearly indicate two annual cycles characterized by an increase and decline in carbon fixation. The lowest standard deviation is found in June, while the highest standard deviation occurs between July and December, probably due to precipitation and temperature variability.

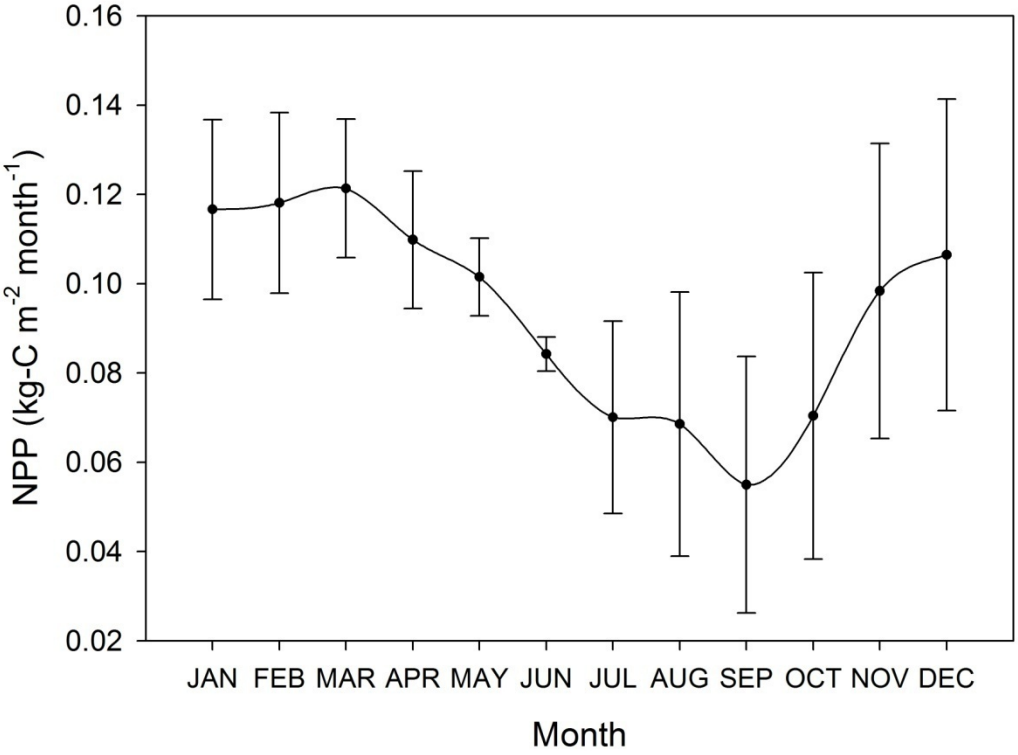

Figure 2.4. Seasonal variation in mean NPP for PDG site (with standard deviation).

There is a pronounced seasonal variation of NPP at the FNS grassland site (Fig. 2.4), with high values of NPP (> 0.05 kg-C m^{-2} $month^{-1}$) from January to May (wet season and wet – dry transition season) and a maximum peak value of NPP in March (0.08 ± 0.014 kg-C m^{-2} $month^{-1}$). The values of NPP begin to fall until September, when it reaches the minimum values (0.004 ± 0.022 kg-C m^{-2} $month^{-1}$), and from this month onwards, NPP increases slowly. Thus, like the cerrado, the grassland shows one cycle of increase and another of decline of carbon fixation. March, October, November and December have the lowest standard deviation, while the highest values are found in January.

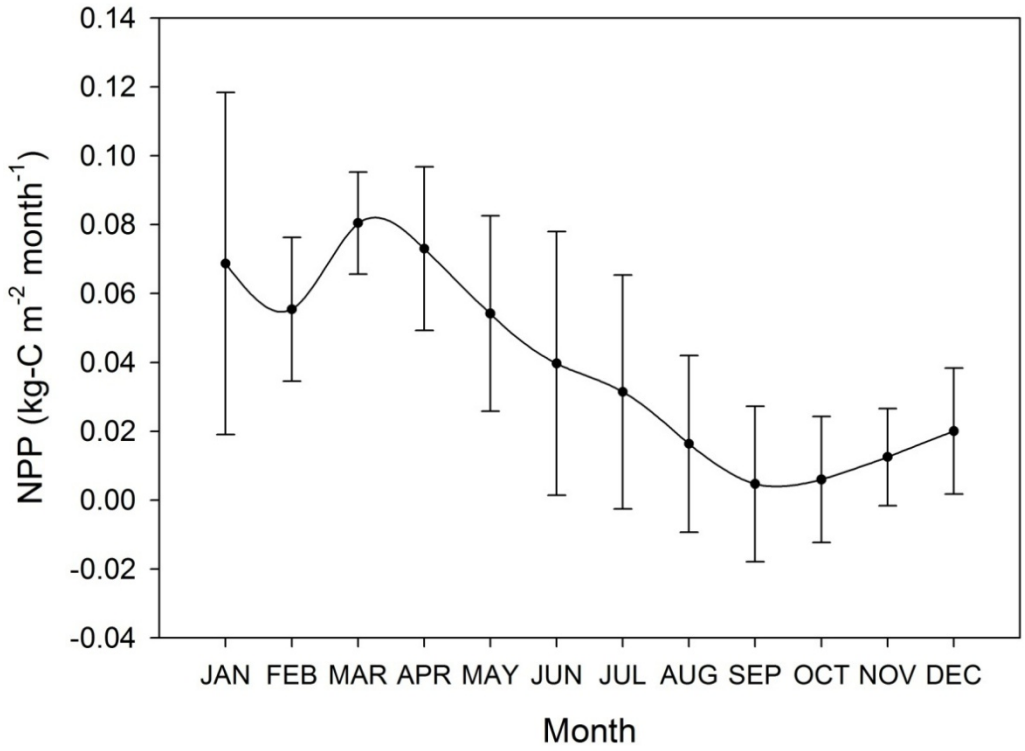

Figure 2.5. Seasonal variation in mean NPP for FNS site (with standard deviation).

Carbon fixation is on average approximately 61% higher in the cerrado than the grassland (Fig. 2.5). Throughout all months the behavior of NPP is similar for both ecosystems, possibly because they are both subject the similar climate variability.

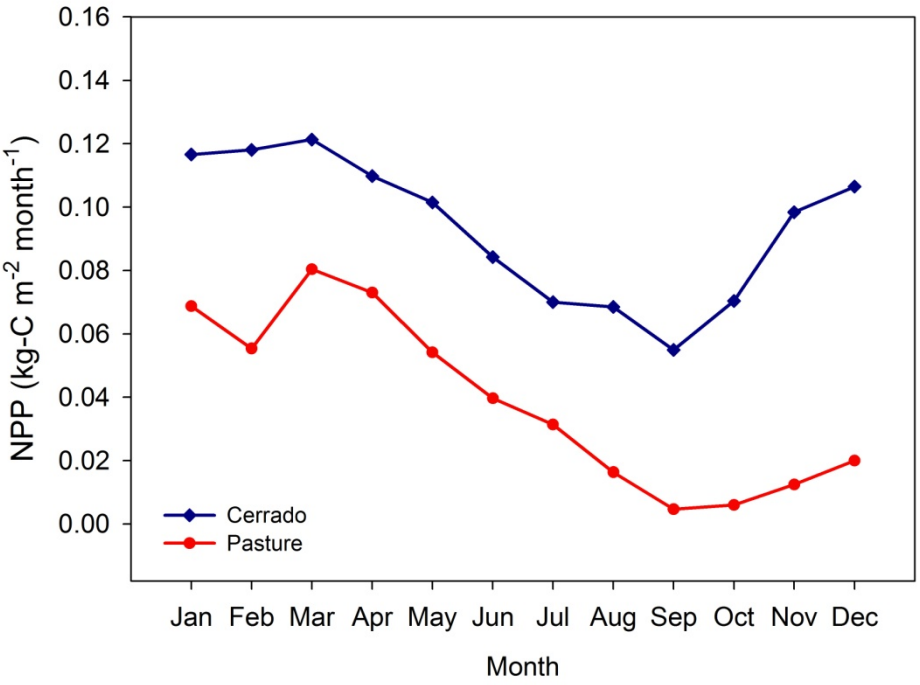

Figure 2.6. Monthly evolution of estimated NPP average for PDG (cerrado site) and FNS (grassland site) sites in 2010.

2.3.2. Annual NPP patterns

Spatial variability in NPP for cerrado and grassland ecosystems were analyzed and compared with NPP estimations from MOD17A3. Examining the variability of NPP for cerrado generated by RATE (Fig. 2.6), a similar pattern is apparent in 2001, 2005 and 2010, with high values of NPP (> 1 kg-C m^{-2} $year^{-1}$) in the region near to tropical rain forest, and also for cerrado in the south and southeast of Brazil, Argentina and Paraguay. Lower values of NPP (< 0.75 kg-C m^{-2} $year^{-1}$) are found in central, central west and northeast Brazil. In general, there is a poor agreement between the magnitude of NPP generated from RATE and MOD17A3. MODIS NPP estimates low values in the cerrado near to tropical rain forest and also for northeast Brazil (< 0.75 kg-C m^{-2} $year^{-1}$). The region with best agreement between the two estimation methods is in the south and southeast of Brazil.

Grace et al. (2006) documented an average NPP of the 0.72 kg-C m^{-2} $year^{-1}$ for tropical savannas around the world and estimated a NPP of 1.06 kg-C m^{-2} $year^{-1}$ for the cerrado ecosystem of Brazil. The authors suggest that the productivity of savannas (or cerrado, in this case) is attributable to the generally lower precipitation, the occurrence of a pronounced and prolonged dry season, or in some cases, the soil characteristics in relation to water-holding capacity.

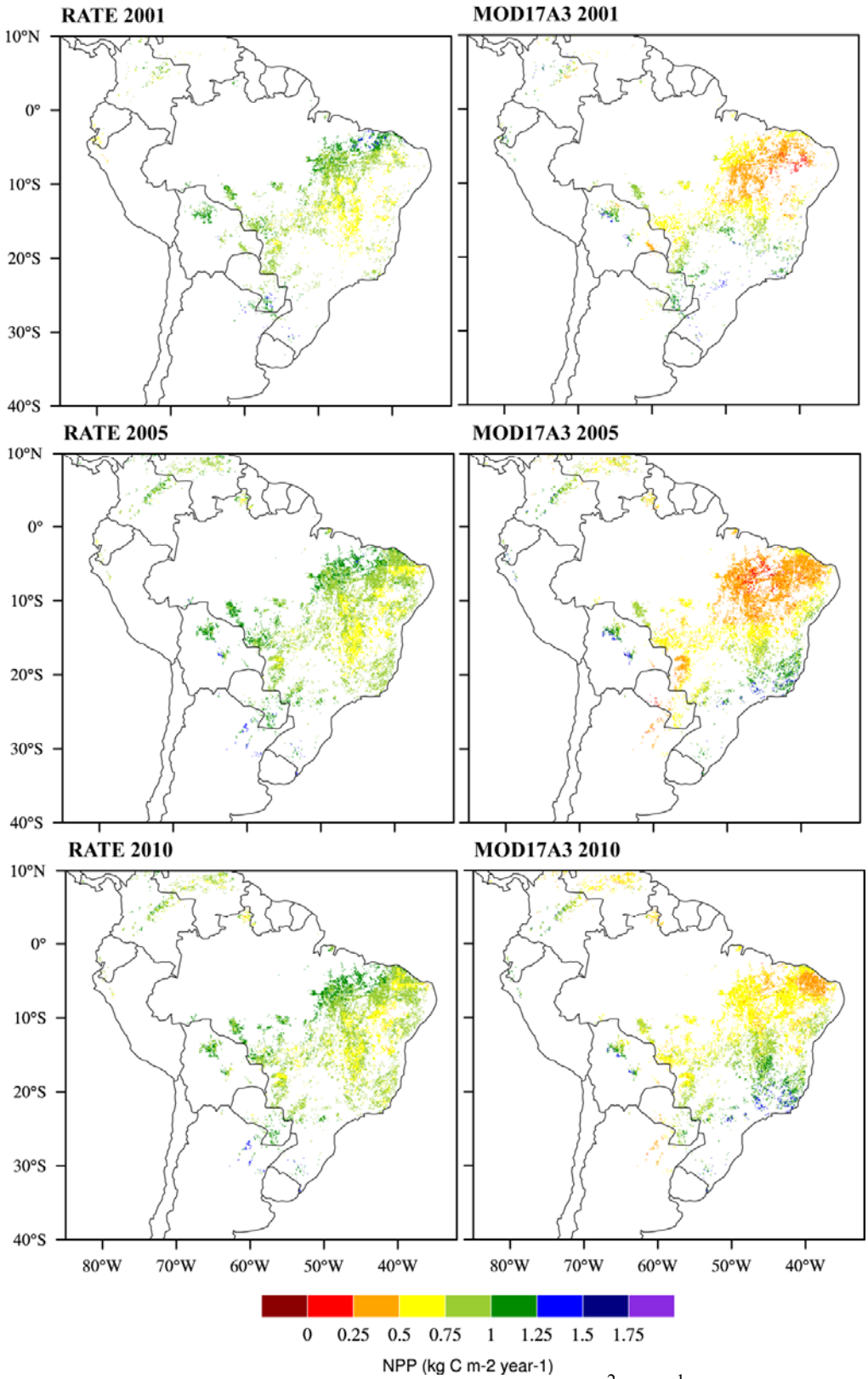

Figure 2.7. Cerrado spatial annual NPP patterns (kg-C m^{-2} $year^{-1}$) in 2001, 2005 and 2010.

The variability of NPP for grassland generated by RATE (Fig. 2.7) shows a similar general pattern in 2001, 2005 and 2010. There are many pixels of grassland in northeast of Brazil showing highly variable values of NPP. More consistent values of NPP are generated for central Brazil, ranging from 0.25 to 0.75 kg-C m^{-2} $year^{-1}$. However, for grassland in southern Brazil, Uruguay and Argentina the values of NPP are higher, ranging from 0.25 to 1 kg-C m^{-2} $year^{-1}$. The grassland of central Argentina, Paraguay and Peru are characterized by a lower value of NPP than other places, probably because it is natural grassland rather than the agriculture grassland predominantly found in south Brazil, Uruguay and other parts of Argentina. Turner et al. (2006) estimated NPP of around of 0.47 kg-C m^{-2} $year^{-1}$ for temperate grassland ecosystems but noted that agricultural grasslands can cause higher values.

A comparison of regional NPP estimated from RATE and MOD17A3 gives, in general, a better agreement for grassland than for cerrado, but is still characterized by considerable differences in simulated values. We calculated Willmott's (1981) the index of agreement (*d*) to assess the level of agreement between punctual NPP estimated from RATE and MOD17A3. In a punctual analysis for PDG and FNS sites, low values of *d* were obtained for cerrado and grassland ecosystems, 0.39 and 0.34 respectively, confirming the general lack of concordance. The large differences between the estimates of NPP from RATE and MOD17A3 are probably due to differences in how NPP is calculated, differences in input data and the specifics of model calibration.

The NPP MODIS product has a tendency to overestimate NPP at low productivity sites and underestimate NPP at high productivity sites (Turner et al., 2006). The overestimation appears to be associated with high estimates of FAPAR

and the underestimation is primarily a function of low values for the maximum light use efficiency.

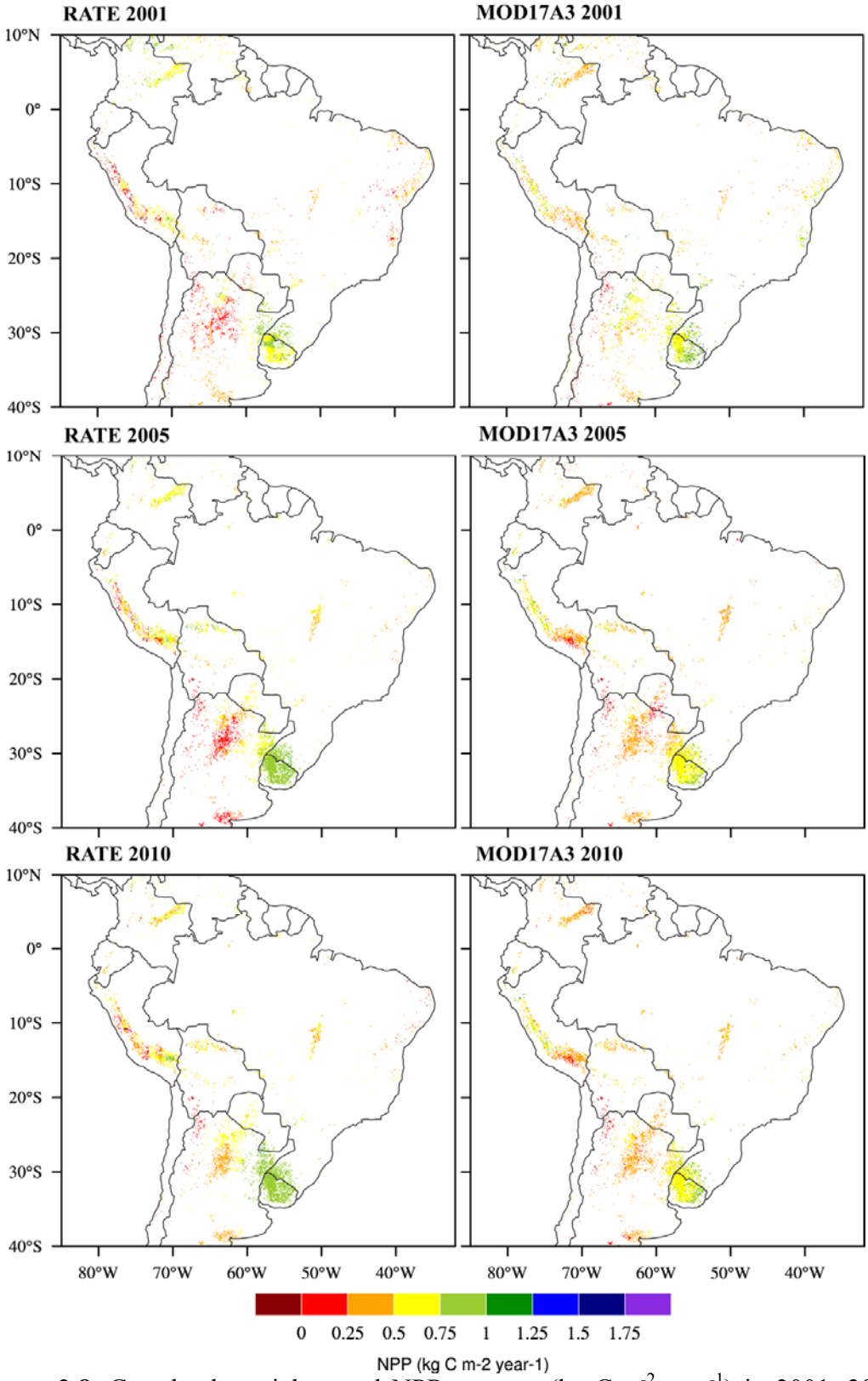

Figure 2.8. Grassland spatial annual NPP patterns (kg-C m^{-2} $year^{-1}$) in 2001, 2005 and 2010.

The PDG and FNS sites of cerrado and grassland both show large interannual variability in NPP, which may be a reflection of strong climatic constraints. General trends in the magnitude of NPP are typically associated with local climate and land use (Turner et al., 2006). The PDG site (Fig. 2.8) generated high values for carbon fixation between 2001 and 2005, peaking in 2004 with a value of 1.31 kg-C m^{-2} $year^{-1}$. In 2006, the NPP had the lowest value among the years analyzed (1.05 kg-C m^{-2} $year^{-1}$). Only 2002, 2003 and 2004 had large standard deviations. The FNS site (Fig. 2.9) had more interannual variability than the cerrado site, with low values of NPP from 2001 to 2004 and 2010 (<0.40 kg-C m^{-2} $year^{-1}$). On the other hand, NPP increased from 2005 to 2009, and had a rapid decline in 2010. In 2009, carbon fixation peaked with a value of 0.74 kg-C m^{-2} $year^{-1}$. Standard deviations were uniformly low except for 2002, 2003 and 2004.

The large variability in NPP estimates for grassland can be explained by the fact that grassland species are sensitive to precipitation variability due to their shallow rooting depths (< 50 cm). These limit water availability to that obtained from atmospheric precipitation and exclude access to groundwater (Shen et al., 2009).

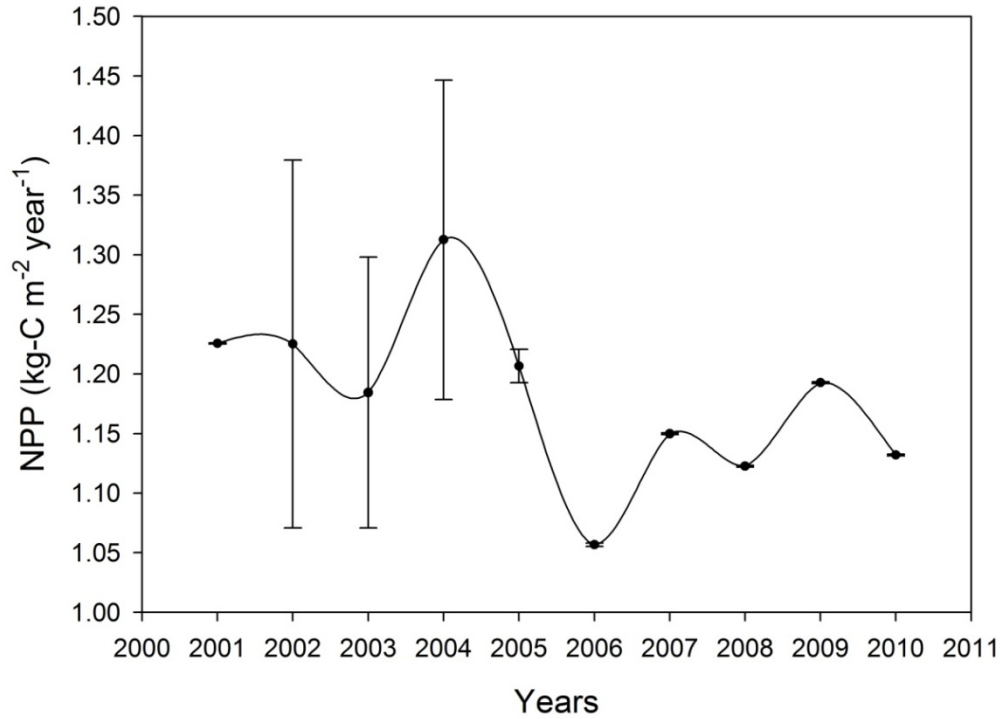

Figure 2.9. Interannual variability in mean NPP for PDG site estimated by RATE (with standard deviation).

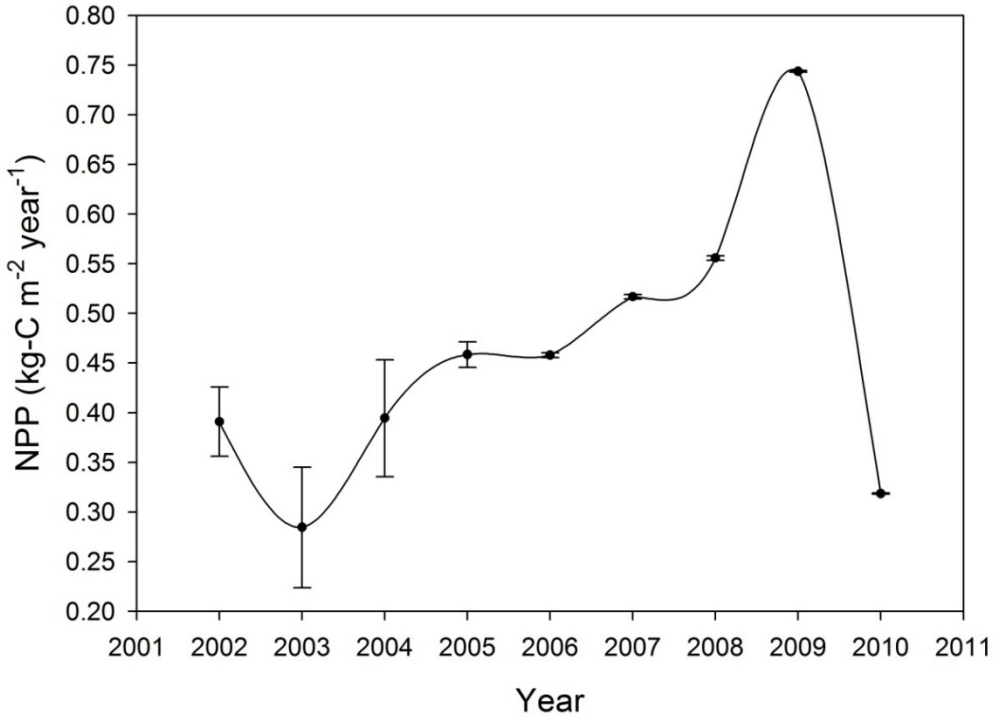

Figure 2.10. Interannual variability in mean NPP for FNS site estimated by RATE (with standard deviation).

2.4. CONCLUSIONS

Net primary production is an important variable for assessing the global carbon cycle. Nonetheless, knowledge about macro-geographic variability in NPP is still relatively limited for many ecosystems. For example, in contrast to tropical forests there are few studies that provide information about NPP for the cerrado and grassland ecosystems of South America. The accurate modeling of NPP in such ecosystems is essential to gain a more complete and accurate understanding of spatial and temporal patterns of NPP at the regional level, especially when empirical data on NPP is sparse or non-existent.

Our monthly NPP estimates for cerrado and grassland in 2010 using the RATE algorithm show a pronounced seasonal variability. Two growing seasons are clearly apparent, one of them characterized by an increase of NPP values, the other by a decrease. These results can be directly linked the observed climate variability of the region, and correspond to periods of high and low precipitation and temperature. Moreover, interannual NPP variability was higher for grassland than cerrado. This may be driven by the increased susceptibility of grassland to variations in climate (especially precipitation) due the short roots of grassland vegetation.

Our results indicate that NPP estimates from variability RATE and the MOD17A3 product are not in good agreement. These differences are probably driven

by differences in how NPP is calculated and the input variables used. Unfortunately, due to the lack of empirical data for NPP in these ecosystems it was not possible to fully validate the estimations of RATE Algorithm for either cerrado or grassland ecosystems. Such validation should be a priority for future research and highlights the great need for basic research in this area.

GENERAL CONCLUSIONS

Studies of the variability and the magnitude of carbon assimilation are indispensable for an improved understanding of carbon dynamics in terrestrial ecosystems. This is especially true to ecosystems such as the cerrado and agricultural pasture that have received very little attention in studies of carbon cycling. Given the lack of observational NPP data, a better understanding of the variability in NPP can only be achieved through the development of models that accurately capture the key ecosystem processes. Nonetheless, these models need to be accurately calibrated if they are to provide realistic simulations.

The results of the study described in Chapter 1 indicate that multi-site calibration is possible, and can produce consistent results across sites - with some restrictions regarding the choice of objective function. The optimal choice of objective function depends on the intended use of the model: On a short time scale, we recommend the site-weighted method using mean absolute error as adjustment measure. For longer times scales we recommend using the site-weighted maximum bias error objective function.

The results described in Chapter 2 indicate a pronounced seasonal variability in monthly NPP estimates for cerrado and grassland in 2010 which is linked to the climate variability of the region - periods of high and low precipitation

and temperature driving seasonal differences in carbon fixation. Interannual NPP variability was observed to be higher for grassland than cerrado, perhaps indicating that grassland may be more susceptible to climate variability. There were also clear differences between MODIS NPP estimates and those of RATE. These differences are probably driven by disparities in how NPP is calculated and the input variables used. Such results indicate that further studies are urgently needed to rigorously validate these alternative strategies to estimate NPP.

More generally, future research needs to extend the approaches developed here to provide a more in depth and systematic analysis of the relative strengths, weaknesses and uses of multi-site calibration for land surface models. For example, an assessment of multi-site calibration for tropical forests and other important ecosystems and a evaluation of multi-site calibration using more than two sites. Another important lacuna in current knowledge is due to the lack of experimental research to provide empirical data about NPP. This is still vital to improve validation of models and to generate better carbon assimilation estimations for ecosystems for which empirical data still do not exist.

GENERAL REFERENCES

ALTON, P.B. How useful are plant functional types in global simulations of the carbon, water, and energy cycles? **Journal Geophysical Research,** v. 116, G01030, doi:10.1029/2010JG001430, 2011.

BEKELE, E.G.; NICKLOW, J.W. Multi-objective automatic calibration of SWAT

using NSGA-II. **Journal of Hydrology,** v. 341, 3-4, p. 165-176, 2007.

BONAN, G.B. A Land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: technical description and user's guide. Technical Report. **National Center of Atmospheric Research**, Boulder, CO, NCAR Technical Note/TN- 417+STR, p. 150, 1996.

BRASWELL, B.H.; SACKS, W.J.; LINDER, E.; SCHIMEL, D.S. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. **Global Change Biology,** v. 11, p. 335–355, DOI: 10.1111/j.1365-2486.2005.00897.x, 2005.

CANADELL, J.G.; LE QUÉRÉ, C.; RAUPACH, M.R.; FIELD, C.B.; BUITENHUIS, E.T.; CIAIS, P.; CONWAY, T.J.; GILLETT, N.P.; HOUGHTON, R.A.; MARLAND, G. Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks. **Proceedings of the National Academy of Sciences,** v. 104, p. 18866-18870, doi: 10.1073/0702737104, 2007.

CAO, W.; BOWDEN, W.B.; DAVIE, T.; FENEMOR, A. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. **Hydrological Processes,** v. 20, p. 1057–1073, doi: 10.1002/hyp.5933, 2006.

COHEN, W.B.; MAIERSPERGER, T.K.; TURNER, D.P.; RITTS, W.D.; PFLUGMACHER, D.; KENNEDY, R.E.; KIRSCHBAUM, A.; RUNNING, S.W.; COSTA, M.H.; GOWER; S.T. MODIS land cover and LAI collection 4 product quality across nine sites in the western hemisphere. **IEEE Transactions on Geoscience and Remote Sensing,** v. 44, 7, p. 1843-1857, 2006.

COSTA, M.H.; PIRES, G.F. Effects of Amazon and Central Brazil deforestation scenarios on the duration of the dry season in the arc of deforestation. **International Journal of Climatology,** v. 30, p. 1970–1979, DOI: 10.1002/joc.2048, 2010.

CRAMER, W.; KICKLIGHTER, D.W.; BONDEAU, A.; MOORE, B.; CHURKINA, G.; NEMRY, B.; RUIMY, A.; SCHLOSS, A.L. Comparing global models of terrestrial net primary productivity (NPP): overview and key results and the participants of the Potsdam NPP Model intercomparison. **Global Change Biology,** v. 5 (suppl. 1), p. 1–15, 1999.

D'HEYGERE, T.; GOETHALS, P.L.M.; de PAUW, N. Genetic algorithms for optimisation of predictive ecosystems models based on decision trees and neural networks. **Ecological Modelling,** v. 195, p. 20–29, doi: 10.1016/j.ecolmodel.2005.11.005, 2006.

da ROCHA, H.R.; MANZI, A.O.; CABRAL, O.M.; MILLER, S.D.; GOULDEN, M.L.; SALESKA, S.R.; COUPE, N.R.; WOFSY, S.C.; BORMA, L.S.; ARTAXO, P.; VOURLITIS, G.; NOGUEIRA, J.S.; CARDOSO, F.L.; NOBRE, A.D.; KRUIJT, B.; FREITAS, H.C.; VON RANDOW, C.; AGUIAR, R.G.; MAIA, J.F. Patterns of water and heat flux across a biome gradient from tropical forest to savanna in Brazil. **Journal of Geophysical Research,** v. 114, G00B12, doi:10.1029/2007JG000640, 2009.

DENMAN, K.L.; BRASSEUR, G.; CIAIS, C.A.; COX, P.M.; DICKINSON, R.E.; HAUGLUSTAINE, D.; HEINZE, C.; HOLLAND, E.; JACOB, D.; LOHMANN, U.; RAMACHANDRAN, S.; SILVA DIAS, P.L., WOFSY , S.C.; ZHANG, X. Couplings Between Changes in the Climate System and Biogeochemistry. **In: 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, p. 499-587, 2007.

FENSHOLT, R.; SANDHOLT, I.; RASMUSSEN, M.S.; STISEN, S.; DIOUF, A. Evaluation of satellite based primary production modelling in the semi-arid Sahel. **Remote Sensing of the Environment,** v. 105, p. 173-188, 2006.

FIELD, C.B.; RANDERSON, J.R.; MALMSTRÖM, C.M. Global net primary production: combining ecology and remote sensing. **Remote Sensing of the Environment**, v. 51, p. 74-88, 1995.

FOLEY, J.A.; LEVIS, S.; COSTA, M.H.; CRAMER, W.; POLLARD, D. Incorporating dynamic vegetation cover within global climate models. **Ecological Applications,** v. 10, p. 1620–1632, 2000.

FOLEY, J.A.; PRENTICE, I.C.; RAMANKUTTY, N.; LEVIS, S.; POLLARD, D.; SITCH, S.; HAXELTINE, A. An integrated biosphere model of land surface processes. **Global Biogeochemical Cycles**, v. 10, p. 603–628, 1996.

FORSTER, P.; RAMASWAMY, V.; ARTAXO, P.; BERNTSEN, T.; BETTS, R.; FAHEY, D.W.; HAYWOOD, J.; LEAN, J.; LOWE, D.C.; MYHRE, G.; NGANGA, J.; PRINN, R.; RAGA, G.; SCHULZ, M.; DORLAND, R.V. Changes in Atmospheric Constituents and in Radiative Forcing. **In: 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, p. 129-234, 2007.

FRIEND, A.D.; ARNETH, A.; KIANG, N.Y.; LOMAS, M.; OGÉE, J.; RÖDENBECK, C.; RUNNING, S.W.; SANTAREN, J.D.; SITCH, S.; VIOVY, N.; WOODWARD, F.I.; ZAEHLE, S. FLUXNET and modelling the global carbon cycle. **Global Change Biology,** v. 13, 610–633, doi: 10.1111/j.1365-2486.2006.01223.x, 2007.

FU, M.C.; GLOVER, F.W.; APRIL, J. Simulation optimization: A review, new developments, and applications. **Proceedings of the 2005 Winter Simulation Conference**, editado por M. Kuhl, N. Steiger, F. Armstrong, e J. Joines, pp. 83–95, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, 2005.

GRACE, J.; SAN JOSE, J.; MEIR, P.; MIRANDA, H. S.; MONTES, R. A. Productivity and carbon fluxes of tropical savannas. **Journal of Biogeography,** v. 33, p. 387–400, 2006.

GRANT, R.F.; HUTYRA, L.R.; de OLIVEIRA, R.C.; MUNGER, J.W.; SALESKA, S.R.; WOFSY, S.C. Modeling the carbon balance of Amazonian rain forests: resolving ecological controls on net ecosystem productivity. **Ecological Monographs,** v. 79(3), p. 445–463, 2009.

GROENENDIJK, M.; DOLMAN, A.J.; VAN DER MOLEN, M.K.; LEUNING, R.; ARNETH, A.; DELPIERRE, N.; GASHA, J.H.C.; LINDROTH, A.; RICHARDSON, A.D.; VERBEECK, H.; WOHLFAHRT, G. Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. **Agricultural and Forest Meteorology,** v. 151, p. 22–38, 2011.

HODNETT, M. G.; OYAMA, M. D.; TOMASELLA, J. Comparisons of long-term soil water storage behaviour under pasture and forest in three areas of Amazonia. In: Gash, J. H. C.; Nobre, C. A.; Roberts, J. M.; Victoria, R. L. **Amazonian Deforestation and Climate**. Chichester: John Wiley, 1996. p. 57-78.

HUANG, N.; NIU, Z.; WU, C.; TAPPERT M.C. Modeling net primary production of a fast-growing forest using a light use efficiency model. **Ecological Modelling,** v. 221, p. 2938–2948, 2010.

IMBUZEIRO, H.M.A. **Calibration of the IBIS model in the Amazonian forest using multiple sites.** 2005. 67f. Dissertação (Mestrado em Meteorologia Agrícola), Universidade Federal de Viçosa, Viçosa, MG.

ISE, T.; LITTON, C.M.; GIARDINA, C.P.; ITO, A. Comparison of modeling approaches for carbon partitioning: Impact on estimates of global net primary production and equilibrium biomass of woody vegetation from MODIS GPP. **Journal of Geophysical Research,** v. 115, G04025, doi:10.1029/2010JG001326, 2010.

JU, W.; WANG, S.; YU, G.; ZHOU, Y.; WANG, H. Modeling the impact of drought on canopy carbon and water fluxes through parameter optimization using an ensemble Kalman filter. **Biogeosciences Discuss,** v. 6, p. 8279–8309, 2009.

KNORR, W.; KATTGE, J. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. **Global Change Biology,** v. 11, p. 1333–1351, doi: 10.1111/j.1365-2486.2005.00977.x, 2005.

KUCHARIK, C.J.; FOLEY, J.A.; DELIRE, C.; FISHER, V.A.; COE, M.T.; LENTERS, J.D.; YOUNG-MOLLING, C.; NORMAN, J.M.; RAMANKUTTY, N. Testing the performance of a dynamic global ecosystem model: water balance, carbon balance end vegetation structure. **Global Biogeochemical Cycles,** v. 14, p. 795–825, 2000.

LEGATES, D.R.; MCCABE JR., G.J. Evaluating the use of ”goodness-of- fit” measures in hydrologic and hydroclimatic model validation. **Water Resources Research,** v. 35(1), p. 233–241, 1999.

LEITH, H. Primary production: Terrestrial ecosystems, **Human Ecology,** v.1 (4), p. 303–332, 1973.

LIU, Y.; BASTIDAS, L.A.; GUPTA, H.V.; SOROOSHIAN, S. Impacts of a parameterization deficiency on offline and coupled land surface model simulations. **Journal of Hydrometeorology,** v. 4, p. 901–914, 2003.

LUYSSAERT, S.; CIAIS, P.; PIAO, S.L.; SCHULZE, E.D.; M. JUNG, M.; S. ZAEHLE, S.; SCHELHAAS, M.J.; REICHSTEIN, M.; CHURKINA, G.; PAPALE, D.; ABRIL, G.; BEER, C.; GRACE, J.; LOUSTAU, D.; MATTEUCCI, G.; MAGNANI, F.; NABUURS, G.J.; VERBEECK, H.; SULKAVA, M.; VANDERWERF, G.R.; JANSSENS, I.A. The European carbon balance. Part 3: forests. **Global Change Biology,** v. 16, p. 1429–1450, doi: 10.1111/j.1365-2486.2009.02056.x, 2010.

LUYSSAERT, S.; INGLIMA, I.; JUNG, M.; RICHARDSON, A.D.; REICHSTEIN, M.; PAPALE, D.; PIAO, S.L.; et al. CO2 balance of boreal, temperate, and tropical forests derived from a global database**. Global Change Biology,** v. 13, p. 2509-2537, 2007.

MAHADEVAN, P.; WOFSY, S.C.; MATROSS, D.M.; XIAO, X.; DUNN, A.L.; LIN, J.C.; GERBIG, C.; MUNGER, W.; CHOW, V.Y.; GOTTLIEB, E.W. A satellite-based biosphere parameterization for net ecosystem CO_2 exchange:

Vegetation Photosynthesis and Respiration Model (VPRM). **Global Biogeochemical Cycles,** v. 22, doi: 10.1029/2006GB002735, 2008.

MALHI, Y.; ARAGÃO, L.E.O.C.; METCALFE, D.B.; PAIVA, R.; QUESADA, C.A.; ALMEIDA, S.; ANDERSON, L.; BRANDO, P.; CHAMBERS, J.Q.; da COSTA, A.C.L., HUTYRA, L.R.; OLIVEIRA, P.; PATINO, S.; PYLE, E.H.; ROBERTSON, A.L.; TEIXEIRA, L.M. Comprehensive assessment of carbon productivity, allocation and storage in three Amazonian forests. **Global Change Biology,** doi: 10.1111/j.1365-2486.2008.01780.x, 2009.

MATSUSHITA, B.; XU, M.; CHEN, J.; KAMEYAMA, S.; TAMURA, M. Estimation of regional net primary productivity (NPP) using a process-based ecosystem model: How important is the accuracy of climate data? **Ecological Modelling,** v. 178, p. 371–388, 2004.

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.; WITTENBERG, U. Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO_2, climate and land use effects with four process-based ecosystem models. **Global Biogeochemical Cycles,** v. 15, p. 183–206, 2001.

MO, X.G.; CHEN, J.M.; JU, W.M.; BLACK, T.A. Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter. **Ecological Modelling,** v. 217, p. 157–173, 2008.

NAYAK, R.K.; PATEL, N.R.; DADHWAL V.K. Estimation and analysis of terrestrial net primary productivity over India by remote-sensing-driven terrestrial biosphere model. **Environ Monit Assess.,** v.170, p. 195–213, 2010.

NEMANI, R.R.; KEELING, C.D.; HASHIMOTO, H.; JOLLY, W.M.; PIPER, S.C.; TUCKER, C.J.; MYNENI, R.B.; RUNNING, S.W. Climate-driven increases in global terrestrial net primary productivity from 1982 to 1999. **Science,** v. 300, p. 1560–1563, 2003.

NUNES, E.L. **Regional algorithm for monitoring the carbon assimilation by tropical forests of the South America.** 2008. 147f. Tese (Doutorado em Meterorologia Agrícola), Universidade Federal de Viçosa, Viçosa, MG.

PETIT, J.R.; JOUZEL, J.; RAYNAUD, D.; BARKOV, N.I.; BARNOLA, J.M.; BASILE, I.; BENDER, M.; CHAPPELLAZ, J.; DAVISK, D.; DELAYGUE, G.; DELMOTTE, M.; KOTLYAKOV, V.M.; LEGRAND, M.; LIPENKOV, V.Y.; LORIUS, C.; PÉPIN, L.; RITZ, C.; SALTZMANK, E.; STIEVENARD, M. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. **Nature,** v. 399, p. 429-436, 1999.

PIAO, S.; CIAIS, P.; FRIEDLINGSTEIN, P.; PEYLIN, P.; REICHSTEIN, M.; LUYSSAERT, S.; MARGOLIS, H.; FANG, J.; BARR, A.; CHEN, A.; GRELLE, A.; HOLLINGER, D.Y.; LAURILA, T.; LINDROTH, A.; RICHARDSON, A.D.;

VESALA, T. Net carbon dioxide losses of northern ecosystems in response to autumn warming. **Nature,** v. 451, 3, January, 2008, doi:10.1038/nature06444, 2008.

PINKER, R.T.; ZHAO, M.; WANG, H.; WOOD, E.F. Impact of satellite based PAR on estimates of terrestrial net primary productivity. **International Journal of Remote Sensing,** v. 31, No. 19, p. 5221–5237, 2010.

PLUMMER, S. On Validation of the MODIS Gross Primary Production Product. **IEEE Trans. Geosci. Remote Sens.,** v. 44 (7), p. 1936-1938, 2006.

POLLARD, D.; THOMPSON, S.L. The effect of doubling stomatal resistance in a global climate model. **Global and Planetary Change,** v. 10, p. 129–161, 1995.

POTTER, C.S.; RANDERSON, J.T.; FIELD, C.B.; MATSON, P.A.; VITOUSCK, P.M.; MOONEY, H.A.; KLOOSTER, S.A. Terrestrial ecosystem production: a process model based on global satellite and surface data. **Global Biogeochemical Cycles**, v. 7(4), p. 811-841, 1993.

POTTER, C.; KLOOSTER, S.; GENOVESE, V. Carbon emissions from deforestation in the Brazilian Amazon region predicted from satellite data and ecosystem modeling. **Biogeosciences Discussions,** v. 6, p. 3031–3061, 2009.

RAYNER, P.J.; SCHOLZE, M.; KNORR, W.; KAMINSKI, T.; GIERING, R.; WIDMANN, H. Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). **Global Biogeochemical Cycles,** v. 19, GB2026, doi:10.1029/2004GB002254, 2005.

RUNNING, S.W.; BALDOCHI, D.D.; TURNER, D.P.; GOWER, S.T.; BARWIN, P. S.; HIBBARD, K.A. A Global Terrestrial Monitoring Network Integrating Tower Fluxes, Flask Sampling, Ecosystem Modeling and EOS Satellite Data. **Remote Sensing of Environment**, v. 70, p. 108-127, 1999.

RUNNING, S.W.; NEMANI, R.R.; HEINSCH, F.A.; ZHAO, M.; REEVES, M.; HASHIMOTO, H. A Continuious Satellite-Derived Measure of Global Terrestrial Primary Production. **BioScience**, v. 54, n. 6, p. 547-560, 2004.

SABINE, C. L.; FEELY, R.A.; GRUBER, N.; KEY, R.M.; LEE, K.; BULLISTER, J.L.; WANNINKHOF, R.; WONG, C.S.; WALLACE, D.W.R.; TILBROOK, B.; MILLERO, F.J.; PENG, T.H.; KOZYR, A.; ONO, T.; RIOS, A.F. The Oceanic Sink for Anthropogenic CO_2. **Science,** v. 305, p. 367-371, 2004.

SANO, E.E.; ROSA, R.; BRITO, J.L.; FERREIRA JR; L.G. Mapeamento semidetalhado do uso da terra do Bioma Cerrado. **Pesq. Agropec. Bras.,** v. 43, p. 153–156, 2008.

SANTOS, S.N.; COSTA, M.H. A simple tropical ecosystem model of carbon, water and energy fluxes. **Ecological Modelling,** v. 176, p. 291–312, 2004.

SCHIMEL, D.S.; HOUSE, J.I.; HIBBARD, K.A.; BOUSQUET, P.; CIAIS, P.; PEYLIN, P.; BRASWELL, B.H.; APPS, M.J.; BAKER, D.; BONDEAU, A.; CANADELL, J.G.; CHURKINA, G.; CRAMER, W.; DENNING, A.S.; FIELD, C.B.; FRIEDLINGSTEIN, P.; GOODALE, C.; HEIMANN, M.; HOUGHTON,

R.A.; MELILLO, J.M.; MOORE, B.; MURDIYARSO, D.; NOBLE, I.; PACALA, S.W.; PRENTICE, I.C.; RAUPACH, M.R.; RAYNER, P.J.; SCHOLES, R.J.; STEFFEN, W.L.; WIRTH, C. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. **Nature**, v. 414, p. 169– 172, 2001.

SCHLOSS, A.L.; KICKLIGHTER, D.W.; KADUK, J.; WITTENBERG, U.; and the Participants of Potsdam NPP Model Intercomparison, Comparing global models of terrestrial net primary productivity (NPP): comparison of NPP to climate and the normalized difference vegetation index. **Global Change Biology,** v. 5 (Suppl. 1), p. 25–34, 1999.

SELLERS, P.J.; BOUNOUA, L.; COLLATZ, G.J.; RANDALL, D.A.; DAZLICH, D.A.; LOS, S.O.; BERRY, J.A.; FUNG, I.; TUCKER, C.J.; FIELD, C.B.; JENSEN, T.G. Comparison of radiative and physiological effects of atmospheric CO_2 on climate. **Science,** v. 271, p. 1402–1406, 1996.

SENNA, M.C.A.; COSTA, M.H.; SHIMABUKURO, Y.E. Fraction of photosynthetically active radiation absorbed by Amazon tropical forest: A comparison of field measurements, modeling and remote sensing. **Journal of Geophysical Research,** v. 110, G01008, doi:10.1029/2004JG000005, 2005.

SHEN, W.; REYNOLDS, J.F.; HUI, D. Responses of dryland soil respiration and soil carbon pool size to abrupt vs. gradual and individual vs. combined changes in soil temperature, precipitation, and atmospheric [CO2]: a simulation analysis. **Global Change Biology,** v. 15, p. 2274–2294, 2009.

SIEGENTHALER, U.; STOCKER, T.F.; MONNIN, E.; LÜTHI, D.; SCHWANDER, J.; STAUFFER, B.; RAYNAUD, D.; BARNOLA, J.M.; FISCHER, H.; DELMOTTE, V.M.; JOUZEL, J. Stable Carbon Cycle–Climate Relationship During the Late Pleistocene. **Science,** v. 310, p. 1313-1317, 2005.

SITCH, S.; HUNTINGFORD, C.; GEDNEY, N.; LEVY, P. E.; LOMAS, M.; PIAO, S.L.; BETTS, R .; CIAIS, P.; COX, P.; FRIEDLINGSTEIN, P.; JONES, C.D.; PRENTICE, I.C.; WOODWARD, F.I. Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). **Global Change Biology,** v. 14, p. 2015–2039, doi: 10.1111/j.1365-2486.2008.01626.x, 2008.

TRUDINGER, C.M.; RAUPACH, M.R.; RAYNER, P.J.; KATTGE, J.; LIU, Q.; PAK, B.; REICHSTEIN, M.; RENZULLO, L.; RICHARDSON, A.D.; ROXBURGH, S.H.; STYLES, J.; WANG, Y.P.; BRIGGS, P.; BARRETT, D.; NIKOLOVA, S. OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models. **Journal of Geophysical Research,** v. 112, G02027, 2007.

TURNER, D.P.; RITTS, W.D.; COHEN, W.B.; GOWER, S.T.; RUNNING, S.W.; ZHAO, M.; COSTA, M.H.; KIRSCHBAUM, A.A.; HAM, J.M.; SALESKA, S.R.; AHL, D.E. Evaluation of MODIS NPP and GPP products across multiple biomes. **Remote Sensing of Environment**, v. 102, p. 282–292, 2006.

WANG, W.; DUNGAN, J.; HASHIMOTO, H.; MICHAELIS, A.R.; MILESI, C. ICHII, K.; NEMANI, R.R. Diagnosing and assessing uncertainties of terrestrial.ecosystem models in a multimodel ensemble experiment: 1. Primary production. **Global Change Biology**, v. 17, p. 1350–1366, doi: 10.1111/j.1365-2486.2010.02309.x, 2011.

WANG, Y.P.; LEUNING, R.; CLEUGH, H.A.; COPPIN, P.A. Parameter estimation in surface exchange models using nonlinear inversion: How many parameters can we estimate and which measurements are most useful? **Global Change Biology,** v. 7, p. 495– 510, 2001.

WHITE, L.K.; CHAUBEY, I. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. **Journal of the American Water Resources Association,** v. 41(5), p. 1077–1089, 2005.

WILLIAMS, M.; RICHARDSON, A.D.; REICHSTEIN, M.; STOY, P.C.; PEYLIN, P.; VERBEECK, H.; CARVALHAIS, N.; JUNG, M.; HOLLINGER, D.Y.; KATTGE, J.; LEUNING, R.; LUO, Y.; TOMELLERI, E.; TRUDINGER, C.M.; WANG, Y.P. Improving land surface models with FLUXNET data. **Biogeosciences,** v. 6, p. 1341–1359, 2009.

WILLMOTT, C.J. On the validations of models. **Physical Geography,** 2, 2, pp. 184-194, 1981.

WILLMOTT, C.J. Some comments on the evaluation of model performance, **Bulletin American Meteorological Society,** v. 63, p. 1309–1313, 1982.

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research,** v. 30, p. 79–82, 2005.

YUAN, H.; DAI, Y.; XIAO, Z.; JI, D.; SHANGGUAN, W. Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. **Remote Sensing of Environment,** v. 115, p. 1171–1187, 2011.

YUAN, W.; LIU, S.; YU, G.; BONNEFOND, J.M.; CHEN, J.; DAVIS, K.; DESAI, A.R.; GOLDSTEIN, A.H.; GIANELLE, D.; ROSSI, F.; SUYKER, A.E.; VERMA, S.B. Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data. **Remote Sensing of Environment,** v. 114, p. 1416–1431, 2010.

ZAKS, D.P.M.; RAMANKUTTY, N.; BARFORD, C.C.; FOLEY, J.A. From Miami to Madison: Investigating the relationship between climate and terrestrial net primary production. **Global Biogeochemical Cycles,** v. 21, GB3004, doi: 10.1029/2006GB002705, 2007.

ZHANG, X.; SRINIVASAN, R.; VAN LIEW, M. Multi-site calibration of the SWAT model for hydrologic modeling. **Transactions of the ASABE,** v. 51, p. 2039–2049, 2008.

ZHANG, X.; SRINIVASAN, R.; VAN LIEW, M. On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. **Hydrological Processes,** v. 24, p. 955–969, doi: 10.1002/hyp.7528, 2010.

ZHAO, M.; HEINSCH, F.A.; NEMANI, R.R.; RUNNING, S.W. Improvements of the MODIS terrestrial gross and net primary production global data set. **Remote Sensing of Environment,** v. 95, p. 164-176, 2005.

ZHU, L.; CHEN, J.M.; QIN, Q.; LI, J.; WANG, L. Optimization of ecosystem model parameters using spatio-temporal soil moisture information. **Ecological Modelling,** v. 220, p. 2121–2136, 2009.