

Analysis of Leaf Area Index in the
ECMWF land surface scheme and
impact on latent heat and
carbon fluxes:
Application to West Africa

L.Jarlan ^{1,2}, G. Balsamo ¹, S. Lafont ¹⁻³,
A.Beljaars ¹, J.C.Calvet ⁴, E.Mougin ²

Research Department

¹ECMWF, ²IRD, ³Forest Research UK, ⁴CNRM France

Submitted to Journal of Geophysical Research

December 2007

This paper has not been published and should be regarded as an Internal Report from ECMWF.
Permission to quote from it should be obtained from the ECMWF.



European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: library@ecmwf.int

© Copyright 2007

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

Abstract

Within the Observatory of Natural Carbon of the GEOLAND project (GEOLAND, 2004), the land surface model of the European Centre for Medium Range Weather Forecast (hereafter CTESSEL) has been modified to include an interactive vegetation module. Following this improvement, a land data assimilation system has been developed and the performance of the assimilation of Leaf Area Index (LAI) satellite products is presented in this paper. The CYCLOPES and MODIS LAI satellite derived data sets are compared together with the open-loop simulation of the CTESSEL model. The land data assimilation scheme is based on a simplified 2DVAR. It is theoretically evaluated thanks to a twin experiment. The 2DVAR is then tested over an experimental site of the AMMA project located in Mali. Simulations after and before data assimilation are confronted to ground measurements of LAI and above-ground biomass. The MODIS products are finally used to analyse the CTESSEL LAI over a large window covering the West Africa from 2001 to 2005. The data assimilation of LAI leads to a significant improvement of the simulated phenology of the vegetation. In addition, it is shown that the impact of the LAI data assimilation on Latent Heat Fluxes is limited whereas LAI is one of the main driving force for the CO₂ Net Ecosystem Exchanges (NEE). It is concluded that a LAI climatology is sufficient to simulate latent heat fluxes with acceptable errors whereas an interactive LAI coupled to a data assimilation system of satellite derived products is necessary for NEE fluxes.

1. Introduction

The land surface conditions are of primary importance for climate and weather prediction. Both root-zone soil moisture and vegetation conditions play a vital role in the partitioning of water and energy budgets at the soil-vegetation-atmosphere interface through evaporation processes of the uppermost surface soil layer and plant transpiration (Shukla and Mintz, 1982). This partitioning, in turn, partly controls the thermodynamic and moisture content of the lower troposphere, and thus, climate. The quality of short term to seasonal weather predictions has been shown to strongly depend on a good initialization of the soil moisture (Beljaars et al., 1999). Zeng et al. (1999) and Philippon et al. (2001) among others found theoretical evidence of the influence of vegetation on precipitation at inter-seasonal to inter-decadal time scale in the Sahel where the surface-atmosphere feedbacks are known to be large. Apart from partitioning surface energy fluxes, vegetation also governs the natural CO₂ exchanges between the terrestrial surface and the atmosphere and the CO₂ surface fluxes partly modulate the variability of the concentration of atmospheric CO₂ at time scales ranging from minutes to months (Tucker et al., 1986).

Within this context, the land surface models (hereafter LSMs) aiming to define the boundary conditions for GCMs have been recently improved to include an interactive vegetation dynamics (Arora, 2002). Stated differently, the vegetation physiological parameters of interest for the simulation of hydrological processes (Leaf Area Index -LAI- and stomata resistance) are not prescribed and kept constant anymore but evolve interactively with environmental conditions thanks to biophysically-based models of vegetation growth and photosynthesis. This approach, although tending towards a more realistic representation of the land surface processes and their interaction, increases the number of uncertain parameters of LSMs and thus, the uncertainties of their simulations. Reducing these uncertainties by combining, optimally, model prediction with observations is the objective of data assimilation.

The existing land data assimilation projects (NLDAS, GLDAS, ELDAS, CALDAS) do not include interactive vegetation land surface models, which limits the use of remote sensing data. The main focus in LDAS projects has been soil moisture, although there is a need to account for vegetation biomass to monitor the biosphere vegetation-atmosphere CO₂ exchange. The GEOLAND Integrated Project (2004-2006) co-funded by the European Commission (GEOLAND, 2004), aims at addressing European and global

environment issues, based on the use of remote sensing data. The observatory of natural carbon (ONC) in GEOLAND, will provide a pre-operational global carbon accounting system, dealing with the impact of weather and climate variability on ecosystems fluxes, on daily to seasonal and inter-annual time scales. The solution chosen in the GEOLAND/ONC is to merge the LDAS approach and the interactive vegetation models.

Estimation of vegetation characteristics (biomass, LAI) was one of the first quantitative application of remote sensing through empirical relationships established between vegetation indices such as the Normalized Difference Vegetation Index (NDVI) acquired by the Advanced Very High Resolution radiometer (AVHRR) sensor and ground measurements (e.g. Tucker et al., 1985). Although strong potentialities were highlighted, these empirical relationships depend on the local conditions of the region where they were developed (soil background, vegetation structure). Furthermore, these relationships do not take into account the effects of water and nutrients limitations on vegetation functioning (Prince, 1991; Lo Seen et al., 1995). The use of radiative transfer models, simulating the physics of the interaction between the electromagnetic waves and the surface components, associated to inversion methods should be theoretically better suited to retrieve a quantitative information on vegetation (LAI, above ground biomass) from NDVI observations (e.g. Myneni et al., 1997; Kimes et al., 2000). This approach has been chosen by the MODIS team to extract LAI from measured reflectances (Tian et al., 2000). Gu et al. (2006) have used these products to analyze the LAI into the LSM of the Canadian weather forecast system. In contrast with this work, the LSM has no vegetation growth and photosynthesis modules.

This study presents preliminary results of an observation-modelling system for the monitoring of LAI and water and CO₂ fluxes (or Net Ecosystem Exchange -NEE-) at the European Centre for Medium Range Weather Forecast. It is based on the LSM of ECMWF named TESSEL (Van den Hurk et al., 2000), recently modified to include an interactive vegetation (hereafter named CTESSEL for Carbon-*TESSEL*; Lafont et al., 2007) constrained by satellite derived LAI products. After a short presentation of the data, the model and methods, the system is confronted to local observations of LAI, soil moisture and biomass performed over one of the AMMA super site located in the Gourma region of Mali. The system is then applied to a large window covering the West Africa and the impact of LAI data assimilation on Latent and carbon fluxes is evaluated.

2. Data and study region

2.1. The Agoufou site and measurements

The Agoufou (15.3°N, 1.3°W) study site is located within the AMMA meso-scale site (14.5° - 17.5° N, 1° - 2° W) in the Gourma region in Mali (Figure 1). The Gourma region is located entirely within the Sahel bioclimatic zone and extends to the South of the Niger River between Timbuctu and Gao down to the border with Burkina-Faso. This is mainly a pastoral region enclosed by the annual average 500 and 150 mm isohyets. The rain distribution is strictly mono-modal with rainfall starting in June and ending in September with a maximum in August. The rainy season is then followed by a long dry season characterized by the absence of green vegetation apart from some scattered trees and shrubs. Rangeland vegetation is composed of an herbaceous layer and a sparse woody plant population. Herb growth is strongly influenced by the pattern and magnitude of rainfall events and by the soil moisture regime resulting from topography runoff and local soil texture. Annual herbs germinate after the first rains, in June or July, and unless the plants wilt before maturity owing to a lack of rainfall, the senescence coincides approximately with the end of the rainy

season. The Agoufou site ($1 \times 1 \text{ km}^2$) is a typical Sahelian landscape characterized by gently undulating sand dunes. The total tree and shrub cover is about 4.5%, whereas the grass cover may vary from 0 to about 60% depending on soil moisture availability. The soil is coarse grained or sandy (>90%).

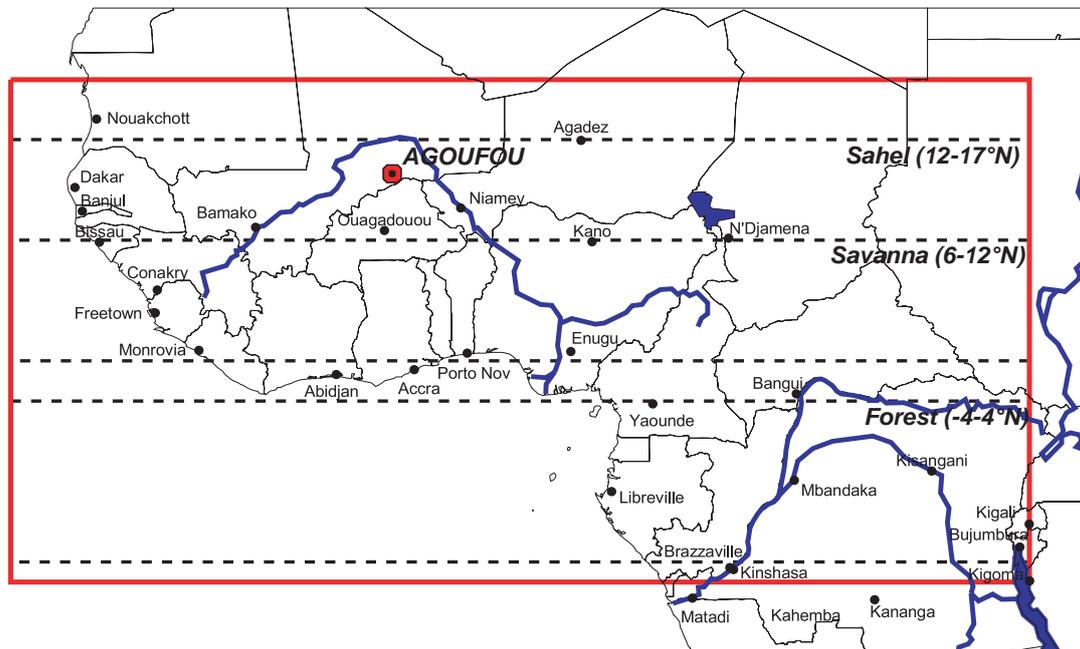


Figure 1: The West African windows (black line). The three sub-region used for the evaluation of the satellite derived data and the CTESSSEL model are represented by dotted line: Sahel ($12-17^{\circ}\text{N}$), Savannah ($6-12^{\circ}\text{N}$) and Forest ($-4-4^{\circ}\text{N}$). The AMMA super site (near the Agoufou village) located in the Gourma region of Mali is superimposed.

For the 2005 wet season, the annual rainfall total reached 410 mm which can be considered as a relatively wet year (the long-term average being 370 mm). Ground measurements of the vegetation consist in an estimate of the time variation of LAI and above ground biomass from trees and grasses using hemispherical photographs (Weiss et al., 2004) and destructive sampling, respectively. For the grass layer, a 1 km transect has been defined in the E-W direction where measurements are performed every 10 m, resulting in 100 pictures. Concerning above ground herbaceous biomass, 12 samples are used. The large quantity of data is sufficient to capture the spatial variability of the grass layer. The computed mean LAI is assumed to be representative at the 1 km^2 scale. The estimated resulting accuracy in terms of LAI is $0.23 \text{ m}^2 \text{ m}^{-2}$ (at 1 s.d.). Further details on the experimental protocol are detailed in Hiernaux (1984). In 2005, the growth of the grass layer started early in June and reached a maximum LAI of $1.8 \text{ m}^2/\text{m}^2$ and 2100 kg of dry matter / hectare by the end of August. In contrast, the LAI of trees estimated from pictures taken of isolated individual stands remains at values lower than 0.2 throughout the year.

2.2. The West Africa window and the meteorological forcing

West Africa gets most of its annual rainfall during the boreal summer months from June to September. This rainy season is associated with the seasonal reversal of the winds in the lowest level of the atmosphere which is called the monsoon. In winter, the wind blows from the cool continent to the warm ocean. Following the sun apparent movement in the course of the year, the continent warms faster than the ocean. This thermal contrast drives the surface pressure contrast between the ocean (high pressure) and the continent (low pressure) and the set up of the monsoon circulation. The resulting precipitation distribution is characterized

by a strong gradient of rainfall amount from north to south and the duration of the wet season range from 1-2 months to the border between Sahara and Sahel to regularly distributed rainfall events along the year to the south. The annual amount is highly erratic with a higher interannual variability to the north of the window. This climate is locally modulated by the topography (Cameroon Mountains, the Adirars to the north).

The study windows cover the West Africa from -5°N to 20°N and from -20°E to 30°W . The West African is sud-divided in four bioclimatic zones (Figure 1): the Sahara desert to the north; the African Sahel which constitutes the transition between the Sahara and the Sudanian savannah and the Guinean forest to the south. The rainfall gradient (annual amount and length of the wet season) from north to south is associated with a gradient of vegetation (types and fraction cover). The Sahel is mostly covered by steppes composed of open annual herbaceous vegetation and of a sparse trees and shrubs layer. The Sudanian region is covered by savannah and dry woodlands. Rain forests are concentrated to the south of the window over the Guinean region.

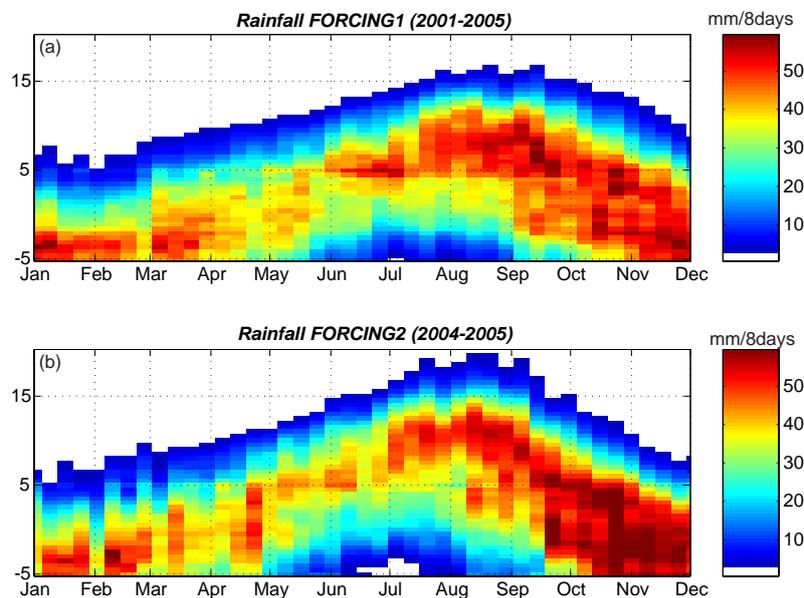


Figure 2: 8 days rainfall Hovmöller diagram for FORCING1 –a- (average over 2001-2005) and FORCING2 –b- (average over 2004-2005).

2.3. The meteorological forcing

The data are provided by the AMMA Land surface Model Intercomparison Project (ALMIP, 2006). ALMIP is part of the AMMA-EU (European Union) and API (Action Programmée Interorganisme: AMMA French program) work packages (WP) 4.1. It is being conducted at CNRM (Centre National de Recherches Météorologiques: National Center for Meteorological Research) and CESBIO (Centre Etudes Spatiales de la Biosphère: Center for the Study of the Biosphere from Space), in Toulouse, France. The data consists in standard meteorological variables provided at a 0.5 degree resolution at a 3 hour time step over the West Africa window from 2001 to 2005. This corresponds to a 101×51 grid which uses a cylindrical equidistant projection. This first forcing data set, named hereafter ‘FORCING1’, is based purely on numerical weather prediction (NWP) forecast model output diagnostics (from ECMWF). Additionally, a second forcing data sets is available for years 2004 and 2005 (hereafter ‘FORCING2’). This dataset is based on the merging of FORCING1 and remote sensing-based products. The downwelling longwave and shortwave radiative fluxes

for 2004 are from the OSI-SAF (Ocean and Sea Ice Satellite Application Facility) and the precipitation is from EPSAT (Estimation of Precipitation by SATellite) product available within AMMA-SAT (AMMA-PRECIP). In 2005, the downwelling radiative fluxes are from LAND-SAF (LAND Satellite Application Facility). Figure 2 displays the Hovmöller diagram of cumulative 8 days rainfall for FORCING1 (Figure 2a) and FORCING2 (Figure 2b). The use of remote sensing-based products based on MSG has resulted in fairly consistent forcing. In particular, it partly corrects for the dry bias of the ECMWF precipitation over Sahel with higher amount than FORCING1 and more rainfall to the extreme north of the window (above 15°N) during July and August. The rainfall amount is also particularly increased from October to December between -5°N and +5°N. A small delay between the two data sets also appeared over Sahel (above 12°N) with the last rainfall events occurring later for FORCING1 than for FORCING2.

2.4. The LAI products

LAI defines an important structural property of a plant canopy as the one sided leaf area per unit ground area. Two LAI products are evaluated and compared to CTESSEL within this study. The first data set, derived from the observations of the VEGETATION instrument, is named CYCLOPES and provided by Medias-France (Toulouse, France) through the core service product of the GEOLAND project. The second one is the extensively used MODIS LAI (Tian et al., 2000).

2.5. The CYCLOPES data set

The VEGETATION instrument on board SPOT4 (launched on April 1998) followed by VEGETATION2 on board SPOT5 (since February 2003) provide with a measure of land surface reflectance in the visible and near infra-red domain continuously. The sensor offers a daily global coverage with a spatial resolution of 1 km². Reflectance measurements are performed within four spectral bands. Further details on the instrument characteristics can be found in Duchemin et al. (2002). The LAI products are processed by the GEOLAND Core Service Products for the years 1998 to 2003. The version 3 of the product is used in this demonstrational system. Version 3 algorithm used re-calibrated and atmospheric corrected reflectances. The LAI is derived from fCover using the semi-empirical approach of Roujean and Lacaze (2002). Further details can be found in the algorithm description (Lacaze, 2004) and in Baret et al. (2007).

2.6. The MODIS data set

The MODIS instrument is operating on both the Terra and Aqua spacecraft. It has a viewing swath width of 2,330 km and covers the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands between 0.405 and 14.385 μm, and it acquires data at three spatial resolutions -- 250m, 500m, and 1,000m. The MOD15 LAI (LAI) is a 1 km global data product updated once every 8-days derived from the MODIS sensor on board TERRA. The MODIS LAI Level 4 algorithms were developed jointly by personnel at Boston University and the University of Montana SCF and NASA GSFC. The algorithm consists of a main procedure based on the inversion of a 3D radiative transfer model thanks to a look-up-table. This algorithm exploits the spectral information content of MODIS surface reflectances at up to 7 spectral bands. Should this main algorithm fail, a back-up algorithm is triggered to estimate LAI and FPAR empirical relationships from vegetation indices. The algorithm uses a land cover classification that is compatible with the radiative transfer model used in their derivation. The LAI products, in the collection 4 version, are available from 2001 to present.

2.7. SIB and CASA NEE

CTESSEL CO₂ fluxes are evaluated against two extensively used data sets (in particular for the TRANSCOM experiment) coming from model simulations: CASA and SIB. The CASA fluxes (Randerson et al., 1997) are based on the Monteith approach: the primary production is the product of the fraction of absorbed radiation (estimated from AVHRR NDVI), the incoming radiation, and an efficiency coefficient. In this simulation the meteorological drivers are based on a climatology. This CASA dataset is currently used as boundary condition by the atmospheric transport model at ECMWF and is expected to be replaced by the CTESSEL fluxes. The SiB 3.0 fluxes (Denning et al., 1996) have been computed with forcing from NCEP2 meteorology data. SiB calculates surface fluxes of sensible and latent heat, radiation, moisture, CO₂, and momentum for vegetated land points. The phenological properties (LAI, vegetation cover) of the vegetation are derived from the GIMMS NDVI data sets (Tucker et al., 2005).

3. The CTESSEL model

The TESSEL model is the land surface model of the ECMWF developed by Van den Hurk et al. (2000). It is designed to describe the exchanges of heat and water between the low-level atmosphere, the vegetation and the soil (i.e. Soil-Vegetation-Atmosphere Transfer -SVAT-) within atmospheric models. Each grid box can be composed of up to 8 tiles: bare soil, high vegetation, low vegetation, high vegetation with snow beneath, snow on low vegetation, interception layer, sea-ice, open water.

During the GEOLAND project, TESSEL has been coupled with the A-gs model (Jacobs, 1994) that described the control of the latent heat fluxes by the plant, diagnoses the LAI and described in a coupled manner the water and carbon cycle. The coupling has been performed by Lafont et al. (2007) and Voogt et al. (2007) based on the work of Météo-France on the land surface model ISBA-A-gs (Calvet et al., 1998). The coupled TESSEL-A-gs is named hereafter CTESSEL for Carbon-CTESSEL. A detailed description of CTESSEL can be found in Lafont et al. (2007) and Voogt et al. (2007) and the main equations of the model are described in appendix 1.

3.1. The ECOCLIMAP data base

At the scale of the West Africa, the ECOCLIMAP data base is used to fill the CTESSEL input parameters. The ECOCLIMAP dataset have been designed to provide a complete set of high resolution surface parameters for land surface models. It is fully described in Masson et al. (2003). The dataset is based on a combination of three products: climate maps, land cover map, and one year of AVHRR NDVI data. A classification process per continent is used to assign a homogeneous ecosystem type (214 in total) to each 1km pixel. All the pixels belonging to a given type have a common land cover, and a similar seasonal cycle. Within this study, the 1km map (214 ecosystems types) ecosystem map is aggregated to the resolution of the meteorological forcing (0.5°) and to a limited number of vegetation types. ECOCLIMAP has been adapted to obtain a 7-type vegetation map:

3 high vegetations types: DECIDUOUS, CONIFEROUS, EVERGREEN

4 low vegetation types: C3 GRASS, C4 GRASS, C3 CROPS and C4 CROPS.

Finally, each grid point can be composed of two kind of vegetation types (low vegetation and high vegetation) with associated fraction cover F_{low} and F_{high} , respectively.

The ECOCLIMAP dataset provides a number of parameters for SVAT models, including albedo, minimum stomatal resistance, roughness lengths and fraction of soil cover. The photosynthesis parameters of the A-gs module (g_c, N_a, e, f ...etc...; see appendix 1) are taken from Gibelin et al. (2006) who describe the implementation at the global scale of ISBA-A-gs and validate the global simulation from satellite observations.

3.1.1. Effect of the forcing and LAI on the simulated water and CO₂ fluxes

In this part, we take benefit of the two forcings to evaluate the impact of the forcing on the simulated LAI, NEE and LE. The 'FORCING2' is characterized by a higher rainfall annual amount than 'FORCING1', particularly over the Sahel (see Figure 2). We will concentrate on this region (between 12°N and 17°N). The CTESSEL model is run for the overlapping period of the two forcing (years 2004 and 2005) with 'FORCING1' and 'FORCING2'. The two experiments are named EXP1 and EXP2, respectively. The average annual cycles of the differences between the two run in terms of LAI, rainfall, NEE and LE are calculated. In addition, two extra runs of the CTESSEL model are done in order to distinguish the impact of LAI on the fluxes (inferred by the different forcings) from the pure impact of the forcing. For this purpose, the LAI is not diagnose anymore but taken from a climatology (ECOCLIMAP). Likewise, for these two additional experiments, the differences between the two runs in terms of rainfall, NEE and LE are calculated. Figure 3 shows these average differences (in percent) over the Sahel from May to October which corresponds to the wet season (where the differences between the two forcings are the most important).

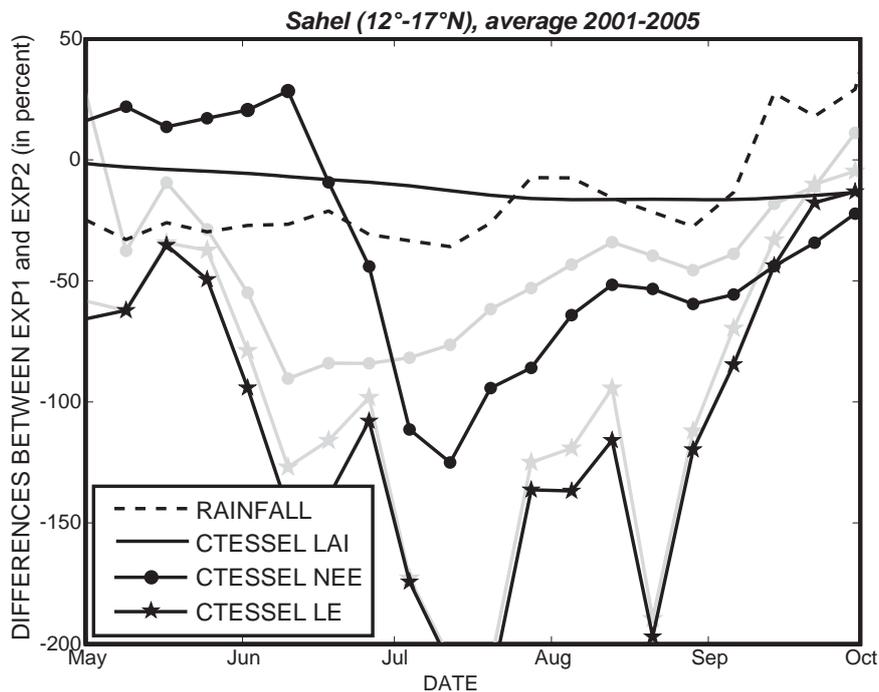


Figure 3: Average cycles of the differences between EXP1 and EXP2 (see text) for RAINFALL, LAI, NEE and LE over Sahel. The grey lines correspond to simulation made with a climatology of LAI (and not a diagnostic LAI).

The rainfall differences between FORCING1 and FORCING2 show the higher rainfall of EXP2 until the beginning of August. FORCING1 maintains more rainfall afterwards as previously underlined. The simulated LAI responds consistently to the higher rainfall provided by FORCING2. In particular, the rainfall

difference of about 30% until mid-July leads to a 10% LAI difference at the end of the growing season. This dampened response can be explained by two factors: (1) the high fraction of bare ground over this region leads to a big loss of water by evaporation that is not available for root extraction (2) the vegetation growth over Sahel is not only driven by the total amount of rainfall but also by the distribution of rainfall events (Le Houérou, 1989). The difference in terms of fluxes is not linearly related to rainfall. The water fluxes differences between EXP1 and EXP2 reach more than 200% at the core of the wet season whereas the NEE differences is about 100%. The two additional experiments with a climatology of LAI (the gray lines at Figure 3) are of particular interest. In terms of LE, the differences EXP1 minus EXP2 with a diagnostic LAI or with a climatology exhibit quite a similar behaviour and follow closely the rainfall differences. This means that the LAI differences inferred by the different forcing have a very low impact on the simulated LE. The behaviour is significantly different for the NEE. With the climatology of LAI (gray line + circle), the differences between EXP1 and EXP2 appear strongly driven by the rainfall differences: the simulated NEE of EXP2 is higher than the one simulated with EXP1 apart at the end of the growing season. When the LAI is diagnose, the LAI through the photosynthesis model and the CO₂ assimilation by the plant governs the differences between the two runs. At the beginning of the growing season, the EXP1 NEE is higher than EXP2 NEE. This can be related to a slightly higher LAI for EXP1 or to more favourable rainfall conditions on the previous month that have allowed the vegetation to grow. Nevertheless, the relative differences (in percent) are high but the absolute values of the NEE (not shown) are low. As the vegetation grow, the assimilation of CO₂ by the plant increases more for EXP2 than EXP1 thanks to the more favourable rainfall and higher LAI. Figure 3 illustrates the strong influence of the rainfall on the simulated LAI, NEE and LE and also shows that while the impact of the LAI is relatively low on the simulated LE, it becomes significant on the simulated NEE.

4. The simplified 2DVAR

The schematic description of the method is displayed at Figure 4. The simplified 2DVAR is a variational method which consists in adjusting the simulations to the observations available within the assimilation window (at observations times) by minimizing a cost function J , with respect to a background information \mathbf{x}^b . The general form of J is given by:

$$\begin{aligned}
 \mathbf{J}(\mathbf{x}) &= (\mathbf{x} - \mathbf{x}^b)^T \frac{\mathbf{B}^{-1}}{2} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y} - H(\mathbf{x}))^T \frac{\mathbf{R}^{-1}}{2} (\mathbf{y} - H(\mathbf{x})) \\
 &= J_b(\mathbf{x}) + J_o(\mathbf{x})
 \end{aligned} \tag{1}$$

The cost function has two terms: the background term $J_b(\mathbf{x})$ which measures the distance between the state vector \mathbf{x} and the a priori state \mathbf{x}^b (weighted by the background error matrix \mathbf{B}), and the observation term $J_o(\mathbf{x})$ which accounts for the distance between the vector of observations during the assimilation window, \mathbf{y} , and the simulations weighted by the observation error matrix \mathbf{R} . The projection of the state vector in the observation space is done through the observation operator $H()$. The minimum of the cost function is given by the classical formulation of the best linear unbiased estimate (with the hypothesis that errors follow a normal distribution and that the linearity hypothesis is fulfilled):

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K} (\mathbf{y} - H(\mathbf{x}^b)) \tag{II}$$

where \mathbf{K} is called the gain and is calculated as follows:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T [\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}]^{-1} \tag{III}$$

The simplified 2D-VAR is an assimilation method developed by Balsamo et al. (2004). The main advantage of the method is to avoid the development of adjoint and linear tangent models to minimize J . It assumes a quasi-linear problem close to the background state \mathbf{x}^b by approximating $H()$ by a one-side finite difference. Stated differently, one additional forecast run with perturbed initial condition is required for each state variable. The linearized observation operator \mathbf{H} is calculated from the difference between the initial background and the perturbed background and from the difference between the observed and model predicted variable at the time of the measurement. As such, the simplified 2DVAR is close to an Extended Kalman Filter from which the propagation of the model error matrix is avoided thanks to the constant \mathbf{B} matrix.

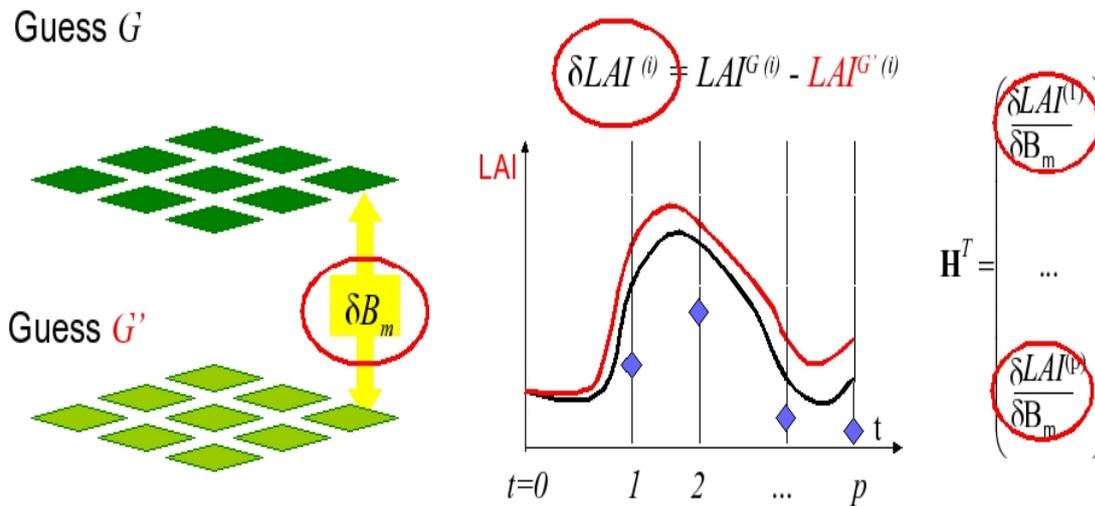


Figure 4: Schematic description of the simplified 2DVAR algorithm to analyse above-ground biomass from LAI observations. A reference and a perturbed run are performed from which the numerical linearization of H is calculated.

The method was initially designed to analyse the root zone soil moisture using 2m air temperature and humidity observations. It has been adapted by Munoz et al. (2007a) to analyse root-zone soil moisture from surface soil moisture observations and to the analysis of both aboveground biomass and root-zone soil moisture by Munoz et al. (2007b). The simplified 2DVAR has been also applied on the SMOSREX data set (Toulouse, France) using point scale ground measurements. The 2DVAR is applied here to the analysis of above ground biomass from satellite derived LAI products over West Africa. The state vector \mathbf{x} is composed of the active above ground-biomass at the grid point scale. For information, it is the same as analyzing LAI (as the ratio between the active above-ground biomass and LAI is constant, see appendix 1). The 2DVAR is applied grid point by grid point. Therefore, \mathbf{B} is diagonal. Finally, the increments $\mathbf{K} (\mathbf{y} - H(\mathbf{x}^b))$ are distributed to the low and high vegetation LAI thanks to their cover fraction F_{low} and F_{high} , respectively.

4.1. The background and observations errors

The two different available forcing datasets are used to evaluate the diagonal terms of the background error covariance matrix **B**. CTESSEL is run with FORCING1 and FORCING2 for the years 2004 and 2005. The average annual cycle of the LAI differences between the two experiments are plotted at Figure 5 in the shape of Hovmöller diagram.

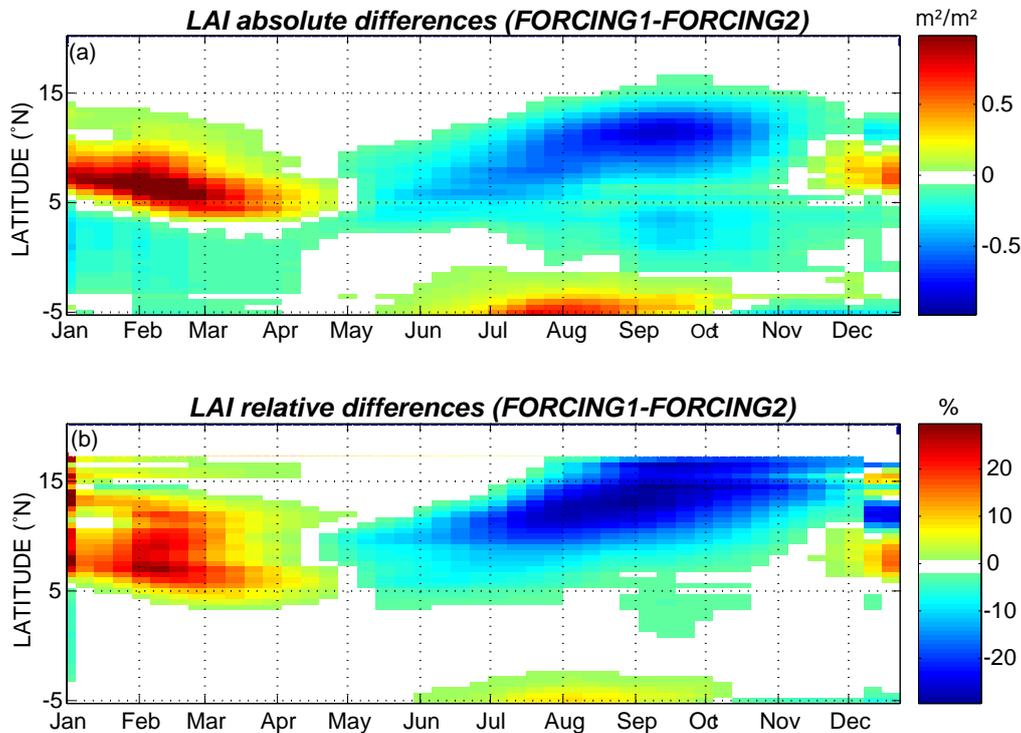


Figure 5: Hovmöller diagram of the LAI differences between CTESSEL forced by 'FORCING1' and CTESSEL forced by 'FORCING2' averaged over 2004-2005: (a) absolute differences (in m^2/m^2); (b) relative differences (in %).

The analysis of Figure 5 can give us indication of the background error. The differences in terms of precipitation between the two forcings are located between 5 and 15°N (Figure 5). Therefore, we focus on this region where the LAI differences are obviously the larger (Figure 5). The lower precipitation of FORCING1 lead to lower simulated LAI than for the FORCING2 experiment. In addition, these LAI differences are mainly associated to the growth phase of the vegetation during July, August and September. During the first half of the year, 'FORCING1' LAI is higher than 'FORCING2' LAI. This may be attributed to the later rainfall events of FORCING1 than FORCING2 that keep quite high LAI until late in the winter. The absolute and relative differences between the two experiments range from -0.9 to +1.2 m^2/m^2 and -29 to +27%, respectively. From these elements, the background error is set semi-empirically to 20% of forecasted LAI with low and high bounds at 0.2 and 1.2 m^2/m^2 over the whole domain of study. Nevertheless, the use of this rather empirical and spatially homogeneous model errors statistics is a first approximation and better estimations, variable in space and time, should be considered in the future. The analysis-ensemble method (Kucukkaraca and Fisher, 2006) which is based on an ensemble of forecasts seems a promising technique.

The satellite derived LAI uncertainty proceeds from different sources (atmospheric and directional corrections of the reflectance used in the inversion process, inversion errors, spatial representativity ...etc...). Therefore, the errors of satellite derived LAI are difficult to estimate and ecosystem dependent.

Within this study, it is empirically fixed to $1 \text{ m}^2/\text{m}^2$. Finally, errors are assumed to be uncorrelated between two 8/10-days satellite acquisitions.

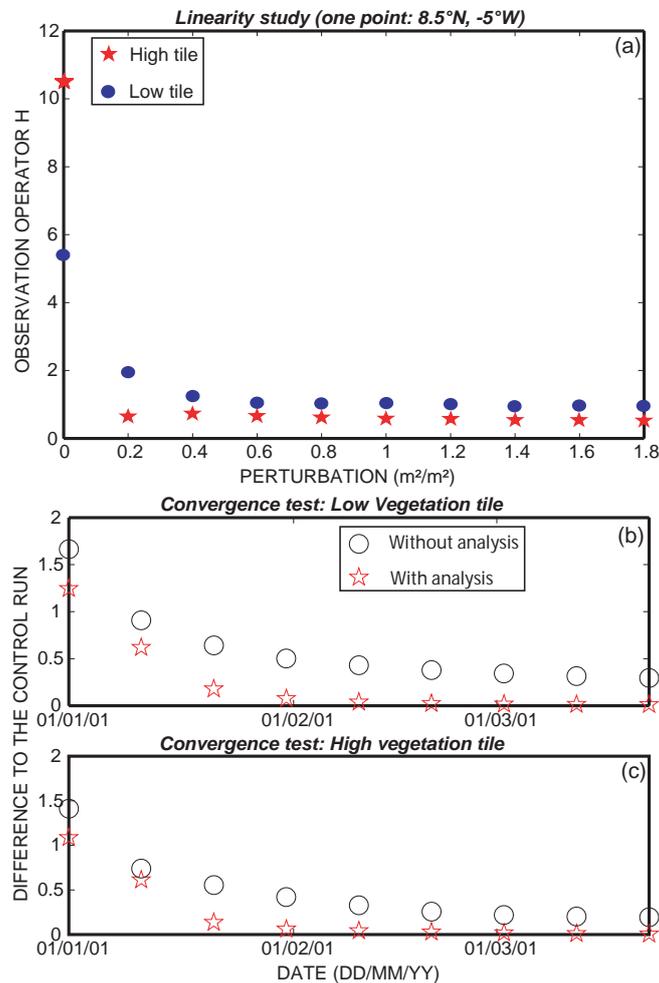


Figure 6(a): Test of the linear hypothesis for the simplified 2DVAR method on 1 point. Apart from numerical errors due to low perturbations, the system is very linear for perturbations chosen above $0.4 \text{ m}^2/\text{m}^2$; (b), (c): Convergence test: absolute difference between the open-loop (circle) and the analysis (stars) and the control run for the low and high vegetation tiles, respectively. The system converges in about one month for the analysis whereas it takes a longer time for the open-loop (without LAI assimilation).

4.2. Testing the simplified 2DVAR

4.2.1. Linearity and choice of the perturbation

The linear approximation used to numerically compute the observation operator has to be evaluated. In order to test the linear hypothesis, an ensemble of 10 perturbations from 0.01 to $1.8 \text{ m}^2/\text{m}^2$ is computed and the observation operator is calculated.

The observation operator is plotted against the size of the perturbation for 1 grid point at Figure 6a. It shows that, above a threshold (around $0.4 \text{ m}^2/\text{m}^2$) the sensitivity of the calculated H doesn't depend on the perturbation value anymore. The perturbation of LAI is taken equal to $0.4 \text{ m}^2/\text{m}^2$.

4.2.2. Evaluation of the analysis convergence

The convergence of the assimilation system is evaluated using twin experiments. A reference run initialized with a reference state of LAI is used to produce a set of LAI simulated “observations” that mimics the satellite temporal repetitivity (10 days). The initial conditions are then modified such as $LAI_{low}=LAI_{high}=3.0$ over the window. A free run (without LAI assimilation) and a run with the analysis of LAI using the artificial observations are computed. Figure 6b and c shows the rmse between these two runs and the control run for the high and the low vegetation tile, respectively. The results show a good convergence of the data assimilation approach after 4 cycles (about 40 days) for both high and low vegetation types. The free run also converges towards the reference run because of the model equation but much more slowly and it still doesn't fit the reference run after 3 months of integration.

5. Evaluation of CTESSEL against the satellite LAI products

The open-loop simulations of CTESSEL are compared to the two satellite derived LAI data sets. The Hovmöller diagrams of the CTESSEL, MODIS and CYCLOPES LAI are presented in Figure 7: (a), (b) and (c), respectively. Only grid points where the satellite products are available are shown.

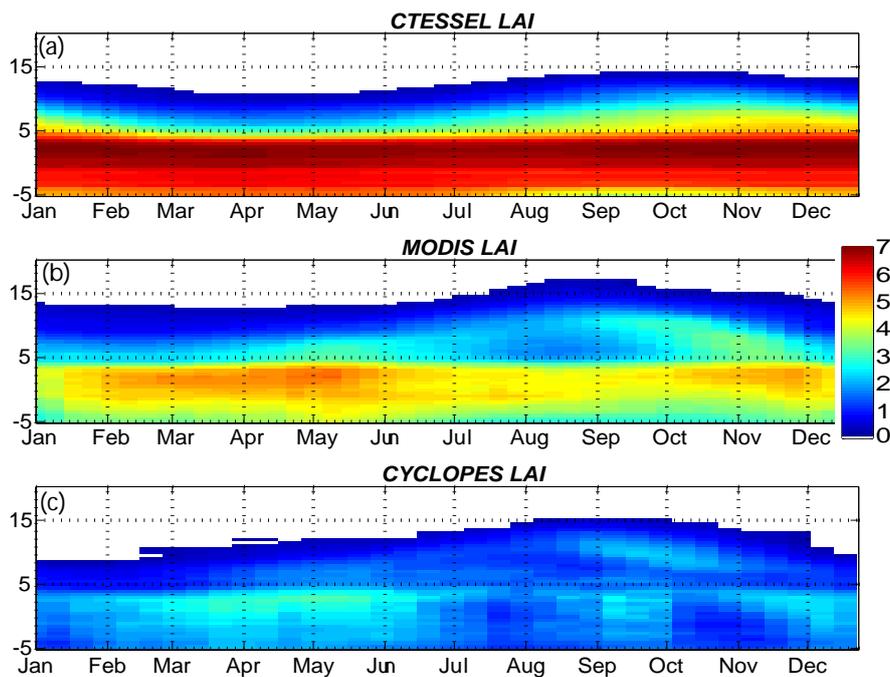


Figure 7: Hovmöller diagram for CTESSEL (a), MODIS (b) and CYCLOPES (c) LAI averaged over 2001-2005 for CTESSEL and MODIS and over 2001-2003 for CYCLOPES.

Above 5°N, the three data sets exhibit a strong seasonality associated to the displacement of the ITCZ to the north during the summer months. The north-south gradient of the length of the growing season is also well marked on the three diagrams. In contrast, the absolute values of LAI are very different. CYCLOPES shows the lower values over the whole region. CTESSEL LAI is clearly the higher below 10°N and MODIS presents intermediate values between the two others. The annual cycles are also differently depicted. Over the so-called Sudanian and Guinean zones, a decrease of LAI centred on August is striking for MODIS and, to a lesser extent for CYCLOPES. This could be attributed to the well known little dry season affecting these

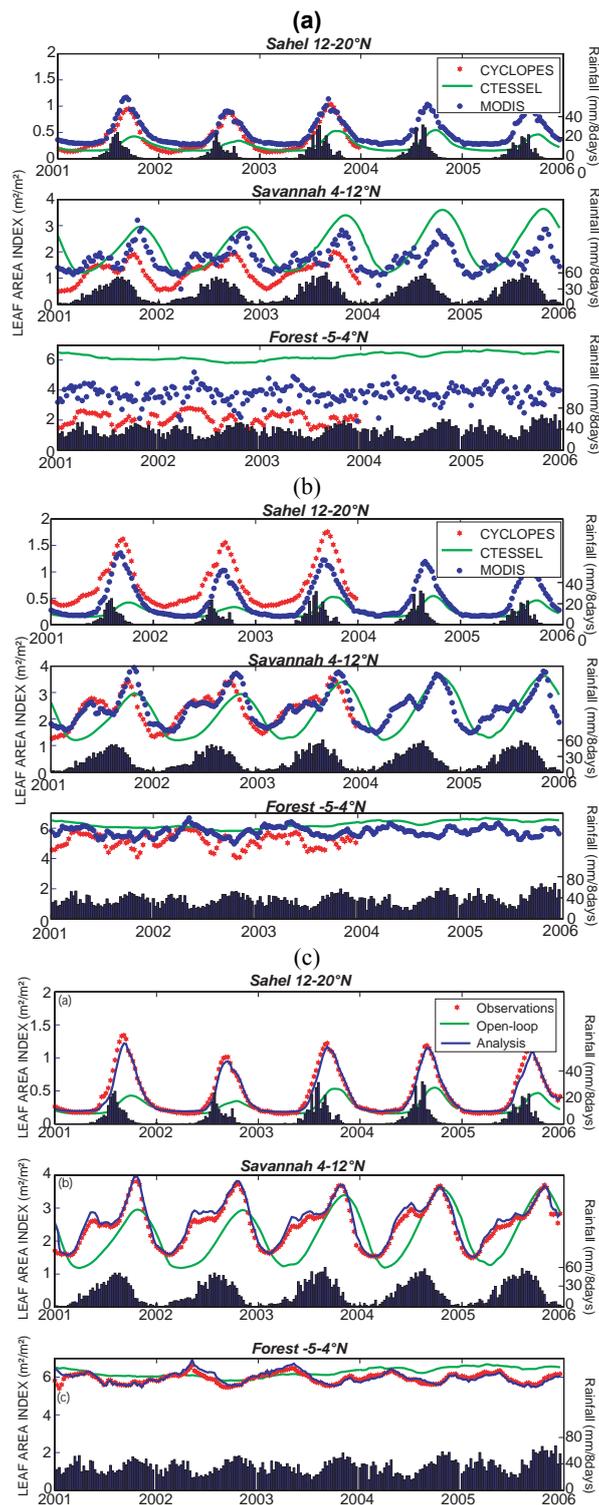


Figure 8: Time series of LAI by latitudinal bands (from top to bottom: Sahel, savannah and forest). (a) Comparison of satellite derived LAI (MODIS and CYCLOPES) before preprocessing (see text) with CTESSEL open-loop; CTESSEL and MODIS are plotted from 2001 to 2005, CYCLOPES is plotted from 2001 to 2003. (b) same as (a) but after rescaling for both satellite data sets and temporal smoothing for MODIS. (c) LAI time series before and after the assimilation of MODIS product into CTESSEL from January 2001 to December 2005. Additionally 8-day rainfall amounts are given by black bars.

areas during August and September but it has been associated to the strong cloud cover and the associated perturbation of the NDVI acquisition by Moulin et al. (1997). The relationship with rainfall (see below, Figure 8) tends to corroborate this assumption. The stable LAI over the Guinean area simulated by CTESSEL is more in agreement with LAI of equatorial forests. Finally, the LAI peak simulated by CTESSEL over Sahel (above 12°N) is lagged by about 2 months with regards to the two satellite data sets. The decrease of LAI after the end of the rainy season is quicker for CYCLOPES than for MODIS.

The time series of LAI for the three data sets are plotted on Figure 8a for three latitudinal bands. Precipitations are superimposed for comparison purposes. Over Sahel, the annual cycle is clearly depicted for the three data sets as stated from the analysis of the Hövmöller. During the dry season, CYCLOPES LAI are lower than MODIS LAI and better match the CTESSEL simulations which is in agreement with the bare soil encountered over the region from October to May. In contrast, the satellite derived LAI are strongly overestimating the CTESSEL simulations. In addition, the peaks of LAI simulated by CTESSEL are lagged of about one month over Sahel with regards to the two satellite products. Over Savannah, they are in phase with MODIS whereas CYCLOPES shows a LAI peak one month earlier. This delay between CTESSEL and the satellite products has also been observed by Gibelin et al. (2006) who evaluate ISBA-a-gs (from which the photosynthesis and growth modules are taken) at the global scale. Over savannah, the decrease of LAI during summer for the MODIS products (and to a lesser extent for CYCLOPES) is not in relation with a significant decrease of rainfall during this period. This decrease is thus attributed to atmospheric perturbations. Over Forest, the overall stable LAI between 6 and 7 m²/m² simulated by CTESSEL is much higher than the satellite products. In particular, the CYCLOPES LAI over forest are far too low (<3 m²/m²). Furthermore, the MODIS data sets appear strongly scattered with regards to CYCLOPES, in particular over savannah and forest. The standard deviations over forest are equal to 0.55 m²/m² and 0.41 m²/m² for MODIS and CYCLOPES, respectively. Finally, the simulated LAI peak value, which generally occurs simultaneously with the maximum evaporation rate and reflects the yearly accumulation of carbon by the vegetation are not in agreement with the satellite products apart from the savannah.

5.1. Preprocessing of the satellite LAI products

5.1.1. Debiasing

The satellite-derived LAI data sets are both negatively biased with regards to the CTESSEL open-loop simulations. The reflectances from which the LAI is retrieved are perturbed and noisy. Because of atmospheric and directional effect, the absolute LAI values are subject to question. The phenology (start, end and peak of the growing season) is assumed to be better depicted by the satellite. In addition, the simplified 2DVAR, similarly to others DA methods derived from the Best Linear Unbiased Estimates, aims at correcting for gaussian errors with a mean equal to zero and not for systematic errors. Therefore, the satellite LAI products are debiased by rescaling their histograms to the CTESSEL LAI simulations by adjusting the 5 and 95 percentile. As such, there is a risk that part of the removed bias originates from the model and not from the observations. This kind of debiasing would tend to reinforce the potential model bias. To limit this risk, the rescaling is computed by vegetation types (thanks to the simplified ECOCLIMAP vegetation map) independently of the geographical area and of the time as one can expect that the bias is localized either temporally (during the summer month where the cloud cover is important for instance) or spatially (over area where the vegetation is dense). This approach ensures that the phenology is kept and that only the absolute values are adjusted. Finally, this approach has shown to give better results than by taking all the LAI values independently of their vegetation types (not shown). For each grid point, the dominant vegetation type is

determined and the rescaling is applied by vegetation types over the study window. For information, four dominant vegetation types are present: evergreen, grass C3, grass C4 and crops C3. The histograms of CYCLOPES and MODIS LAI products (for all vegetation types) are compared to CTESSEL simulated values before and after rescaling at Figure 9.

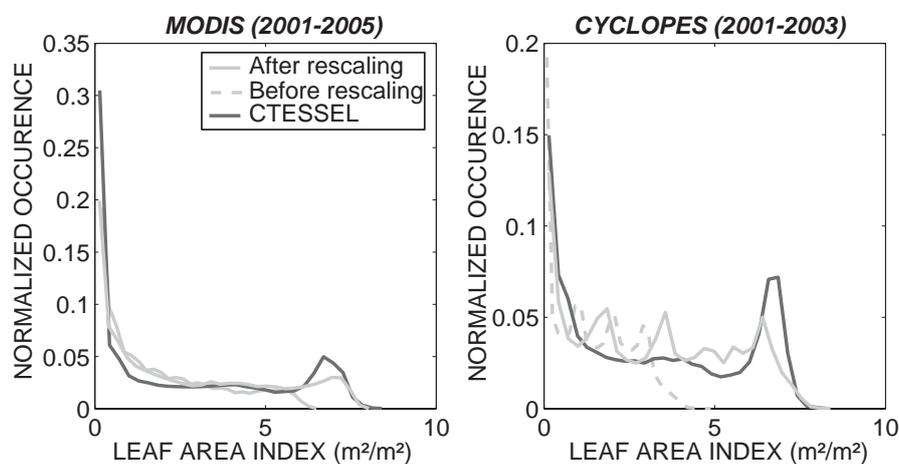


Figure 9: LAI histograms for CYCLOPES and MODIS before and after rescaling to CTESSEL LAI values. The availability of LAI products is 2001-2003 and 2001-2005 for CYCLOPES and MODIS, respectively.

The rescaling is really efficient for MODIS with a smooth curves over the range of LAI values apart from the higher LAI corresponding to equatorial forest that are not totally retrieved after rescaling. In contrast, the histogram of CYCLOPES LAI is more discontinuous with peaks at 2, 4 and 6 m^2/m^2 which are not reproduced by CTESSEL. After rescaling, the peaks remain and the rescaled histogram doesn't match the CTESSEL one properly.

5.1.2. Smoothing

The MODIS data set is available over the whole study period (2001-2005) and has LAI values in better agreement with CTESSEL after rescaling. Nevertheless, it shows strong scatter (see time series on Figure 8a). The second step in the preprocessing is to apply a temporal smoother to MODIS LAI products. The CYCLOPES LAI already presents a smooth temporal evolution. This corroborates the findings of Weiss et al. (2007).

The evolution of LAI has a relatively smooth dynamic and the model is not supposed to match the high frequency oscillation of the satellite LAI products that are attributed to atmospheric perturbations, erroneous clouds masking and differences in the acquisition geometrical configuration. Since these perturbations lead to a negative bias, a upper envelope temporal smoother is applied to the time series over each grid point independently. The Best Index Slope Extraction filter is chosen (BISE filter; Viovy et al., 1992). Since this filter is difficult to parameterize, a preliminary study aiming to test every combination of parameters values is carried out in order to find the best tradeoff between smoothing and preservation of peak values. With this optimal parameter sets, the signal is clear and smooth over most regions. Figure 8b shows the time series of satellite products and CTESSEL LAI (same as Figure 8a) after rescaling and smoothing of MODIS LAIs.

The time series of satellite products are more in agreement with CTESSEL. Over forest in particular, the overall stable LAI of about $6\text{m}^2/\text{m}^2$ matches the model prediction. The satellite products are still affected by a low seasonal signal that is more to be attributed to atmospheric perturbations than to real vegetation

phenology. The scattering of MODIS LAI has been strongly minimized thanks to the smoothing. Over savannah, the two satellite data sets are in good agreement all together and matches the CTESSEL LAI values in terms of the minimum and maximum but the annual cycle, as already highlighted above, strongly differs. Over Sahel, the amplitude of the annual signal is much higher for the satellite products than for CTESSEL. This low annual cycle of CTESSEL is also to be attributed to the dry bias of the precipitation from FORCING1. In addition, because of the rescaling, the dry season LAI value for CYCLOPES is higher than the simulated CTESSEL LAI.

6. Evaluation of the simplified 2DVAR approach on the Agoufou site

The 2DVAR is evaluated on the Agoufou site during the growing season 2005. Ground LAI measurements, extracted every ten days to mimics satellite repetitivity, are assimilated into CTESSEL in order to analyse above ground biomass. The tree cover fraction being low, trees are not considered in this study and CTESSEL is run considering only C4 grasses. Furthermore, the implementation of the data assimilation system with C3 grasses, C3 crops or C4 crops doesn't significantly modify the conclusions of the study. Figure 10 displays the evaluation of the 2DVAR over the Agoufou site: LAI, aboveground biomass, mid-day Latent Heat Fluxes and mid-day NEE averaged over 10 days-period are displayed at Figure 10a, b, c and d, respectively.

This site, located in the Sahelian zone *sensu stricto* as defined by Le Hou  rou (1989) is characterized by a short rainy season of about three months. The vegetation growth, composed mainly of annual herbs, is very rapid and it is followed by a strong senescence after the last rainfall events. This is well depicted by the LAI and above ground biomass measurements (stars on Figure 10a and b). The open-loop simulations of CTESSEL don't match this short growing period. This is in accordance with the findings of the comparison between CTESSEL LAI and the satellite LAI products (over Sahel the simulated LAI peak is delayed with regards to the satellite products). Over the Agoufou site, the CTESSEL peak is also delayed by about 25 days together with the senescence phase that extents further after the last rainfall events during the dry season. In contrast, the value at LAI peak is quite well simulated by CTESSEL (1.45 m²/m² and 1.80 m²/m² for CTESSEL and ground measurements, respectively). This differs from the comparison with the satellite products but, over Agoufou, the meteorological forcing is measured on site whereas over West Africa, the FORCING1 based on the ECMWF forecast is used. The ECMWF operational forecasting system is known to underestimate the sahelian precipitation. The measured annual rainfall is 410 mm/year whereas the corresponding grid point value for FORCING1 is 47 mm/year (against 286 mm/year for FORCING2). Concerning biomass, the 2DVAR data assimilation system also corrects the delay of CTESSEL open loop but the decrease phase appears slightly in advance.

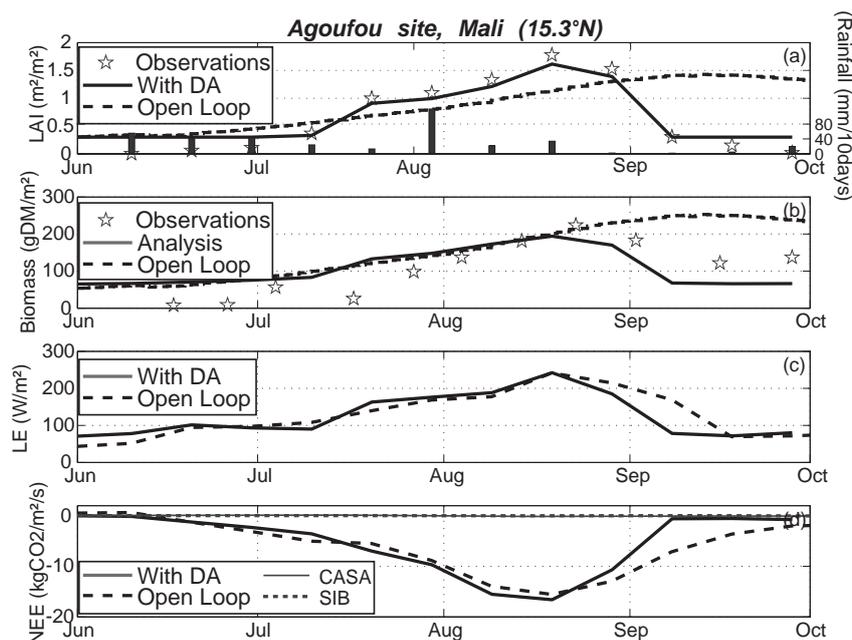


Figure 10: 2D VAR analysis (black line) of above ground biomass from LAI observations over the Agoufou site (Mali) during the year 2005: (a) Leaf Area Index; (b) Above-ground biomass; (c) 10 days average Latent Heat Flux at mid-day; (d) 10 days average Net Ecosystem Exchange at mid-day (SIB and CASA NEE of the corresponding grid point are superimposed). Convention for fluxes is 'positive upward'. The open loop CTESSEL simulations (dotted line) and 10-days precipitations (bar) are superimposed for comparison purposes.

The analysis of above ground biomass and LAI has a low impact on water and NEE during the growing phase (see Figure 10c and 10d). In particular, the peak value of Latent Heat flux (LE) is very close with and without data assimilation -DA- (-242 W/m^2 and -241 W/m^2 , respectively). Sivakumar (1990) underlines that the primary factor governing evaporation variability over Sahel is the rainfall pattern. The impact of the 2DVAR on the NEE peak is also limited (-15.45 and $-16.59 \mu\text{molCO}_2/\text{m}^2/\text{s}$ with and without DA, respectively). During the senescence phase, differences are more striking. The CTESSEL open-loop maintains a high evaporation rate and a high NEE even after the last rainfall events. This can be attributed to the higher LAI and consequently, the higher water extraction from depth soil by the roots. To that respect, the CTESSEL with the LAI analysis has a behaviour better in agreement with what we know of the LE dynamic over Sahel. Wallace and Holwill (1997) among others show a return to very low evaporation rate just 2 days after the last rainfall events thanks to LE measurements performed at the end of the rainy season over a Sahelian site similar to the Agoufou site. For information, SIB and CASA NEE has been superimposed on Figure 10d. Both show almost no variability with fluxes around 0 all around the growing season. These two data sets forced by AVHRR NDVI for vegetation characteristics may suffer from the limitation of this sensor in case of low vegetation cover. CTESSEL NEE may represent a strong improvement over the Sahelian region with regards to these two data sets.

For lack of available LE and CO₂ fluxes on the site, the order of magnitude of CTESSEL simulations is just roughly verified by comparison to measurements already performed by other teams over similar ecosystems. In terms of LE, the daily value (averaged over a 10 days period) reaches -90 W/m^2 (not shown) that corresponds to around $3 \text{ mm H}_2\text{O}/\text{day}$. This is in good agreement with value usually given for evapotranspiration over Sahel during the core of the rainy season such as the measurements performed during the Hapex-Sahel Experiment (Kabat et al., 1997). Concerning NEE, it shows a clear diurnal patterns

(not shown) with the highest uptake rates around noon and a permanent slight release to the atmosphere at night. This is corroborated by measurements of Hanan et al. (1998) performed over a natural savannah. They give noon peak values reaching, on certain days, $-15 \mu\text{molCO}_2/\text{m}^2/\text{s}$ (for LAI peak equal to $0.9 \text{ m}^2/\text{m}^2$). Falk et al. (2007) measured on a site of natural savannah located in Burkina an average noon values of $-22 \mu\text{molCO}_2/\text{m}^2/\text{s}$ on average over July and August (for LAI peak equal to $3.2 \text{ m}^2/\text{m}^2$, C. Brüemmer, pers. Com.). CTESSEL with LAI analysis gives $-16.6 \mu\text{molCO}_2/\text{m}^2/\text{s}$ (average over a 10 days period) (for a simulated LAI peak of $1.45 \text{ m}^2/\text{m}^2$; cf Figure 10).

Concerning daily values, Hanan et al. (1998), gives values up to $\mu\text{molCO}_2 \text{ m}^2/\text{s}$ whereas CTESSEL with LAI analysis is a significantly higher value of $-4.54 \mu\text{molCO}_2 \text{ m}^2/\text{s}$ (not shown) as it is an average over a 10 days period. This may be attributed to the soil respiration which is parameterized from the surface temperature without taking into account soil moisture. Hanan et al. (1998) values of night-time respiratory CO_2 fluxes are comprised between 2 and $5 \mu\text{molCO}_2 \text{ m}^2/\text{s}$ (average value around $3 \mu\text{molCO}_2 \text{ m}^2/\text{s}$). In contrast, CTESSEL (with or without LAI analysis) night-time CO_2 fluxes are around $2 \mu\text{molCO}_2 \text{ m}^2/\text{s}$ which is quite low. In particular, Veenendaal et al. (2004) underline the occurrence of marked spikes of CO_2 release because of rainfall events over a semi-arid savannah in southern Africa. This behaviour is also observed over the Agoufou site after the analysis of the first measurements of NEE performed at the beginning of the growing season 2006 (V. Le Dantec, CESBIO, Toulouse, pers. com.). This suggests the need for a better parameterization of the soil respiration into CTESSEL in the near future. Finally, the return to very low value of the simulated NEE after the last rainfall events has already been observed by Verhoef et al. (1996).

Two additional experiments have been carried out to investigate the effect of negatively biased forcing precipitation (such as FORCING1) on the simplified 2DVAR system. Table 1 displays the peak values of LAI, LE and NEE for the open-loop simulations and the 2DVAR system. Three cases are considered: the initial precipitation, precipitation divide by 2 and precipitation divided by 4. In contrast to the results with the initial forcing precipitation, LAI, LE and NEE peak values strongly differs between 2DVAR and open-loop simulations by introducing a dry bias on the precipitation forcing. In particular, if the precipitation is divided by 4, the 2DVAR system is able to maintain an LAI close to the one simulated with the initial forcing. Furthermore, the LE and NEE are divided by 2 whereas they are divided by 3 and about 4 for the open-loop, respectively. As a conclusion, the impact of the analysis of LAI on LE and NEE is low if the water availability is sufficient for vegetation growth (apart from the delayed decrease of the fluxes after the last rainfall). When the precipitation is negatively biased, the analysis of LAI allows maintaining a significantly higher evaporation rate and NEE than with the open-loop. This is of particular interest for the simulation of the African Monsoon by the NWP model of ECMWF.

		LAI (m^2/m^2)	LE (W/m^2)	NEE ($\mu\text{molCO}_2/\text{m}^2/\text{s}$)
PP	Open-loop	1.43	241.6	-15.45
	2DVAR	1.62	242.3	-16.59
PP/2	Open-loop	1.09	152.2	-7.95
	2DVAR	1.62	180.1	-9.09
PP/4	Open-loop	0.7	81.2	-4.31
	2DVAR	1.62	117.2	-7.94

Table 1: Peak values of LAI, LE and NEE for CTESSEL open-loop and CTESSEL analysis for different value of precipitation (measured precipitation -PP-, measured precipitation divided by two -PP/2- and measured precipitation divided by four -PP/4-).

7. Application to West Africa

The 2DVAR algorithm has been evaluated using twin experiments and over a site where ground measurements were available. The objective of this part is to apply this algorithm to the real MODIS LAI products (after rescaling and smoothing). The impact of the LAI data assimilation on the simulated Latent heat and NEE is evaluated. Results are compared to the TESSEL model (without the interactive vegetation part) in terms of LE for information.

7.1. LAI time series

Figure 8c displays the LAI time series for the three latitudinal bands described above. Thanks to the 2DVAR system, the low values of LAI over Sahel are corrected with peak values ranging from 1.0 to 1.5 m²/m² instead of values lower than 0.5 m²/m² for CTESSEL open-loop. Over savannah, the analysis follows very closely the MODIS products apart from the spring months where analyses are slightly above the satellite LAI. Indeed, during the forward integration of the model after the analysis, LAI is decreasing (not shown). This is attributed to the simulated environmental conditions (and in particular the root-zone soil moisture) that are not sufficient to maintain such a high LAI. Around the LAI peak, analysis and open-loop are more in agreement. Finally, the senescence phase is closely following the last rainfall events of the year for the analysis whereas it is too slow for the open-loop (this corroborates the remark made over the Agoufou site above). Another striking feature is also the artificial break in the growing phase of the LAI for the MODIS product during the spring months. It is certainly an artefact of the satellite products due to bad atmospheric corrections. Nevertheless, it has been minimized thanks to the smoothing filter with the upper envelope. Over Forest, the LAI are quite stable for the two LAI apart from noise attributed to atmospheric perturbations of the MODIS products.

7.1.1. Quality of the analysis

For diagnosing the properties of data assimilation systems, we use three quantities defined below:

The differences ($\mathbf{y} - H(\mathbf{x}^b)$) in equation (II) are called **innovations**.

The **residuals**, computed by replacing \mathbf{x}^b by \mathbf{x}^a in the differences above, represent the fit of the observations to the analysis.

Finally, the **analysis increment** reconstructed *a posteriori* by subtracting the background from the analysis.

By lack of ground “truth”, these three quantities can reveal a lot of the performance of the assimilation system. Figure 11 displays the Hovmöller diagram (average over 2001-2005) of the innovations (a), the residuals (b) and the analysis increments(c). Values below 0.2 m²/m² have been masked for a better clarity of the figure.

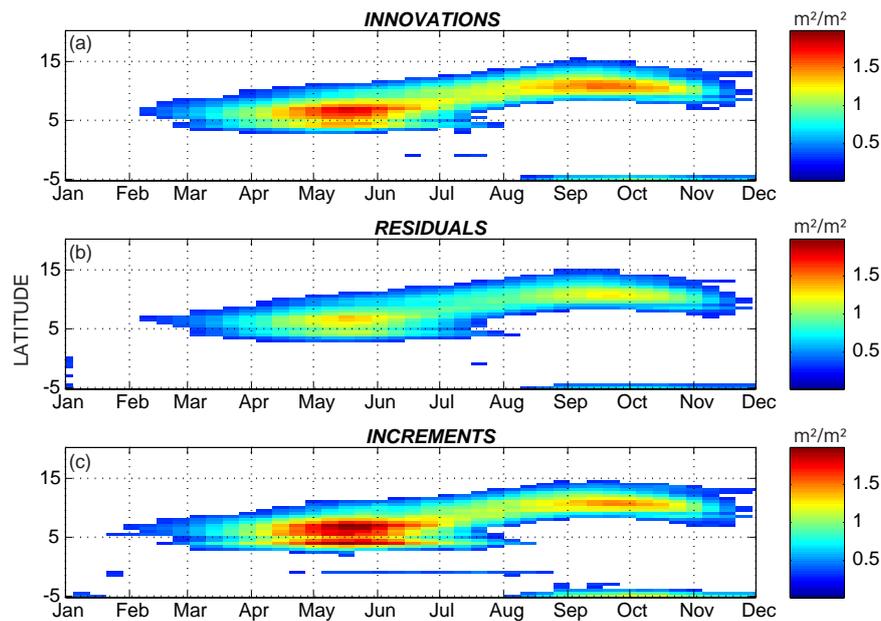


Figure 11: Hovmöller diagram of the innovations (a), the residuals (b) and the analysis increments (c) – see text – averaged over the period 2001-2005.

First, the fit of the observations to the background and analysis can be conveniently examined by Figure 11a and b. As expected, one can see that the differences are smaller for the analysis departures (residuals) than for the background departures (innovations). Stated differently, the LAI analysis draws the CTESSEL simulation to the MODIS products thanks to the simplified 2DVAR. Information has been extracted from the observations. Furthermore, most of the time, the innovations and residuals are low (below $<0.3 \text{ m}^2/\text{m}^2$). Below 5°N , corresponding to equatorial forest areas, this is attributed to the quite constant LAI both for CTESSEL and for the observations (see Figure 8c). Over savannah and Sahel, the residuals and the innovations are also negligible during winter because of the low and constant LAI values corresponding to the dying vegetation of these dry months. The innovations and the residuals are significantly different from zero during spring over savannah and during summer over Sahel. For these two cases, a strong negative bias of CTESSEL has already been identified even after the rescaling process. This means that the analysis is above the observations at the beginning of the assimilation window but, at the time of the observations (around the middle of the assimilation window), it has gone below the observations. Indeed, the environmental conditions are not sufficient to maintain such a high quantity of vegetation and the vegetation is dying during the forward integration of the model.

In a perfect data assimilation system increments should be small and time-average close to zero which would illustrate an absence of bias in the observations. The preprocessing of the data has strongly minimized the bias as shown above (Figure 8b), and the analysis increment (not shown). Nevertheless, some bias still appear over savannah and Sahel. This bias is attributed to (1) the slow growth in the CTESSEL model (e.g. the delayed peak of LAI values in Figure 8b) as well as to (2) the dry bias of the precipitation used to force the model, in particular over Sahel. Nevertheless, it is interesting to point out that even in a case of a biased forcing, the assimilation system is still available to provide with realistic value of LAI. As a conclusion, the system is performing quite well with residuals lower than innovations and analysis increments reasonably small over most of the study window.

Figure 12 displays the average analysis increments over the period 2001-2005. The map (Figure 12a) exhibits a marked contrast between savannah and Sahel to the north with positive increments and Forest on the south where the increments are slightly negative. The histogram (Figure 12) should be centered on zero in the case of unbiased data but it is clearly shifted to positive values. Nevertheless, this shift remains reasonable (mean = 0.05 m²/m²) and it has been strongly reduced thanks to the *a priori* rescaling of the satellite products.

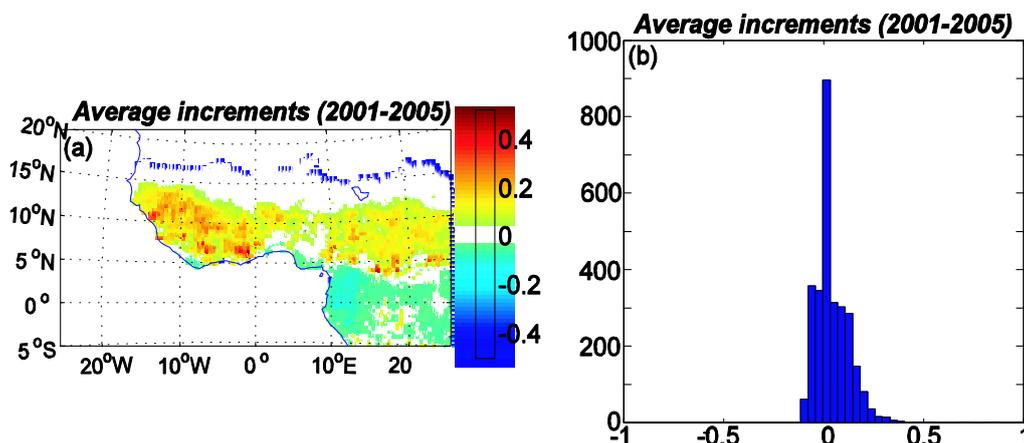


Figure 12: Average analysis increments over 2001-2005: (a) map of West Africa; (b) histogram of values.

7.1.2. Effect of the LAI analysis on simulated Latent and CO₂ fluxes

The objective of introducing an interactive LAI and a data assimilation system in the model of ECMWF is twofold: (1) vegetation, and LAI in particular, affects, the surface latent heat flux (Van den Hurk et al., 2003 among others demonstrate the impact of LAI on the annual land surface evaporation cycle) that is of primary importance for NWP; (2) the coupling of the land surface model with a photosynthesis and a vegetation growth modules is the only way to simulate NEE. Within the frame of its new implication in the Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data (GEMS) project (GEMS, 2005), it is of particular interest for ECMWF to be able to simulate its own CO₂ fluxes consistent with LE to propagate this CO₂ into the atmosphere thanks to the atmospheric dynamical model.

Figure 13 displays the relative differences between CTESSEL+LAI analysis and CTESSEL open-loop for LAI, LE and NEE averaged over 2001-2005. The LAI differences between the two runs are obviously consistent with what has been underlined at Figure 8b. The main LAI differences (on the order of 50% more for CTESSEL analysis) are located over Sahel in summer and over savannah at spring. On average, this 50 % LAI leads to a similar increase in NEE except at the beginning of spring between 7 and 10°N where the increase is higher than 80%. CTESSEL LAI reach their lower values of the year (about 1m²/m², see Figure 8b) whereas the analysis draw the simulations to the MODIS products values, about 1.5 m²/m² higher. The NEE encountered at this time of the year is low (because of a high incoming radiation combined with a high water stress) but this strong gap in LAI leads to significant relative differences in NEE. Likewise, the -30% LAI differences around January over savannah are also reflected in a significant decrease in NEE (on the order of -80%).

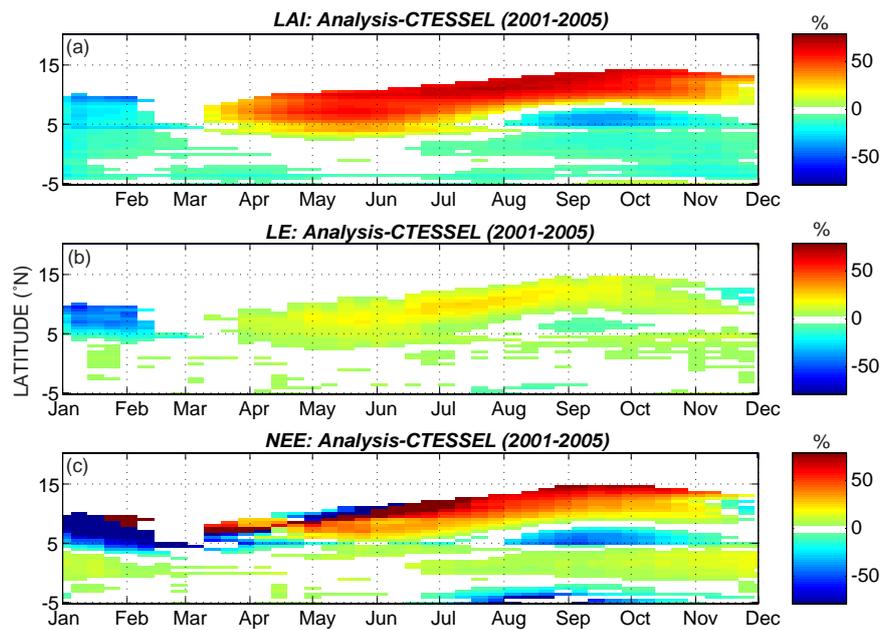


Figure 13: Relative difference between CTESSEL+LAI analysis and CTESSEL open-loop averaged over 2001-2005 with FORCING1 for (a) LAI; (b) LE and (c) NEE. Relative differences ranging from -3% to +3% are masked for clarity purposes.

The impact on LE is much more moderate with a maximum impact of about +24% (corresponding to a LAI difference of +70%). In addition, this maximum impact occurs at the beginning of the rainy season when the absolute values of LE are low. Most of the time, the impact of LAI on LE is lower than 15% (with LAI differences of more than +50%). In addition, the difference reach around -40% over a large area extending between 5°N and 10°N at the end of the rainy season (around January). This is attributed to the shifted senescence phase of CTESSEL open-loop already observed over the Agoufou site (see Figure 10).

In order to investigate the impact of LAI analysis on LE and NEE interannual variability, the normalized standard deviation (hereafter improperly called “variability” for simplicity) over 2001-2005 of LAI, LE and NEE are computed and summarized by Hovmöller diagrams in Figure 14 for Open-loop and CTESSEL+Analysis. Hovmöller of rainfall and TESSEL LE variability are also displayed for comparison purposes on Figure 14g and h, respectively. The variability of 8-day rainfall (Figure 14h) ranges from 22% to 223%. This range of values is much more than for other plotted variables because of the short temporal scale of variability of rainfall. The highest variability is encountered over Sahel. Rainfall, in the Sahel is mostly generated by squall lines, which typically arrive at 3 days intervals throughout the rainy season. Furthermore, storms within these squall lines are convective and the temporal distribution of rainfall is thus highly variable. This high temporal variability combined to low 8-day amount gives high normalized standard deviation (above 100%). Another area of high variability is located to the extreme south of the area during the summer months. During this period of the year, the ITCZ reaches its northern position leading to a significant decrease of the 8-day rainfall amount to the south and consequently, an increase of its variability.

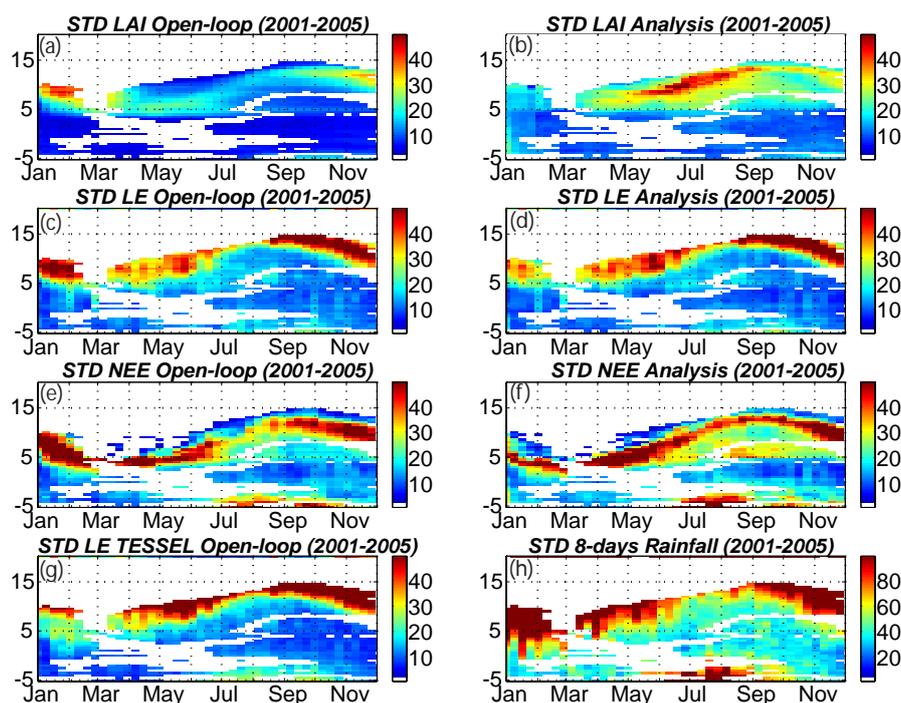


Figure 14: Hovmöller diagram of the (8-days average) normalized standard deviation (2001-2005) of LAI, LE and NEE for CTESSEL open-loop on the left (Figure 14 a,c and e) and for CTESSEL analysis on the right (Figure 14 b, d and f). The standard deviation of TESSEL (constant LAI) over the same period and the standard deviation of the 8 days cumulative rainfall are displayed on Figure 14 g and h for comparison purpose. Areas where LAI analysis and $abs(LAI\ analysis - LAI\ open-loop)$ are below $0.2\ m^2/m^2$ are masked.

The Hovmöller diagrams of LE for CTESSEL open-loop, Analysis and TESSEL (Figure 14c, d and g, respectively) presents quite similar shape. Furthermore, most of the patterns are in good agreement with the ones of precipitations (Figure 14h). In particular, the northern strip of LE high variability from September to January and from April to June can be related to the similar pattern of the 8-day rainfall. These two temporal periods correspond to the beginning and to the end of the rainy season, respectively. During the core of the rainy season, the rainfall events are more regularly distributed (Figure 14h) and LE displays a lower variability on the three runs (CTESSEL open-loop, Figure 14c, CTESSEL analysis, Figure 14d, and TESSEL, Figure 14g). By contrast with the northern area, the area of high rainfall variability to the south during the summer months (already identified above) is not accompanied by a strong LE variability. The high annual rainfall amount allows for a full filling of the soil reservoir. Consequently, even during the short period of lower rainfall, water availability is sufficient to keep high evaporation rate. The day-to-day variability is also further minimized thanks to the high LAI encountered over these region (around $6\ m^2/m^2$, see Figure 8c). In addition, the LAI is not fluctuating a lot (see Figure 14a and b).

In contrast with LE, the LAI and NEE Hovmöller diagrams exhibits more distinct differences between CTESSEL open-loop and CTESSEL analysis. In addition, the Hovmöller patterns are not following so closely the rainfall patterns of variability. The northern strip of rainfall variability doesn't exhibit a high LAI and NEE variability. Over these areas, LAI and the assimilation of carbon by the vegetation (which is mainly governing the NEE) are low and their year to year variability is also below 20%. Just below this northern strip of low variability, LAI and NEE displays their highest variability. This area of high variability on both LAI and NEE is moving northward with time (corresponding to the northward movement of the ITCZ). The

variability of LAI analysis during the core of the rainy season over Sahel and northern savannah (above 10°N) is higher than the one of LAI open-loop. The variability of NEE follows quite closely the variability of LAI with a higher variability of NEE analysis than for the NEE open-loop from May to December. In contrast with LE, the variability of NEE is strong even during the core of the rainy season over Sahel (August/September). This period of the year corresponds to the LAI peak and potentially the highest year-to-year absolute differences in terms of LAI. Nevertheless, the influence of the rainfall on NEE variability is also important. The area of high rainfall variability to the south during the summer months (already identified above) is accompanied by a high variability of NEE (in contrast with LE). Vegetation may strongly decrease the assimilation of carbon during stress period even if the LAI is high. The assimilated carbon will be used to maintain vegetation tissues alive and will not create new tissues. The resulting NEE will decrease. This could explain the highest variability of NEE over these areas

As a conclusion, the analysis of the normalized standard deviation of LAI, LE and NEE shows that the interannual variability of LE is more related to rainfall variability than to LAI whereas NEE variability is in good agreement with LAI variability. This has strong consequences for land surface and, in particular, vegetation modelling for NWP and climate: Figure 13 and Figure 14 shows that a LAI climatology is sufficient to get a reasonable seasonal and interannual variability of LE whereas for NEE, an interactive LAI reacting to environmental conditions combined with a data assimilation system to avoid model divergence is necessary to get realistic variation of NEE at the seasonal and interannual time scales.

7.1.3. Comparison of CTESSEL NEE with SIB and CASA

The CTESSEL NEE with and without LAI data assimilation is compared to two terrestrial surfaces CO₂ flux data set used within the TRANSCOM experiment: the CASA and SIB fluxes already described above. The NEE simulated by CTESSEL open-loop and CTESSEL analysis are plotted together with SIB and CASA simulations at Figure 15 for three latitudinal bands. CASA is a climatology (simulated with average meteorological and NDVI forcings) and SIB has been run for year 2002. The SIB and CASA seasonal cycles are repeated for each year of the study period. Table 2 summarizes the NEE average value over the “growing” season (criterion: NEE<0) and the “dry” season (criterion: NEE>0) for the four data sets.

		CT OL	CT AN	SIB	CASA
dry season NEE>0	Sahel	0.13	0.13	0.21	0.23
	Savan	1.53	1.52	0.72	0.77
	Forest	0.18	0.12	0.56	0.41
growing season NEE<0	Sahel	-0.25	-0.51	-0.21	-0.70
	Savan	-1.45	-1.86	-1.44	-1.07
	Forest	-0.15	-0.25	-0.78	-0.57

Table 2: Average monthly NEE values in $\mu\text{mol CO}_2/\text{m}^2/\text{s}$ by “season” for CTESSEL open-loop (CT OL), CTESSEL with LAI data assimilation (CT AL), SIB and CASA (see text). The dry season is defined as $NEE > 0$ and the wet season by $NEE < 0$.

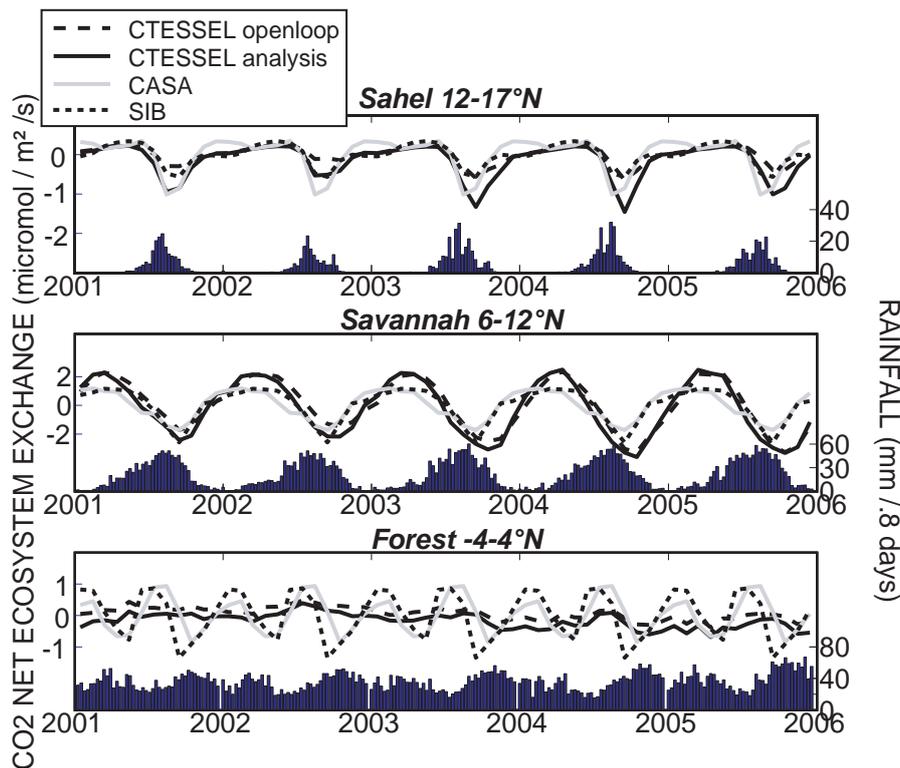


Figure 15: Comparison of monthly NEE by latitudinal bands over 2001-2005 for CTESSEL open-loop and CTESSEL analysis: (a) Sahel; (b) savannah; (c) forest. SIB and CASA simulations (see text) are superimposed for comparison purposes.

Above all, the four NEE data sets are of the same order of magnitude. Over Sahel, CTESSEL with LAI data assimilation leads to higher NEE during the growing season than CTESSEL open-loop because the LAI is increased by the analysis (cf. Figure 8c), and consequently is the assimilation of CO₂ by the plants. During the dry season, CTESSEL open-loop and analysis match together because NEE is mainly governed by the soil respiration that is parameterized by a function of the surface temperature in CTESSEL. CASA and SIB respiration take into account soil moisture that may explain the differences observed with CTESSEL. Notably, at the beginning of the growing season, the first rainfall events are known to cause CO₂ drop into the atmosphere because of soil respiration (Le Dantec, pers. com.). This behaviour seems to be caught by CASA and SIB whereas CTESSEL exhibits a smooth increasing shape from December to June to be associated with the regular increase of the incoming solar radiation (not shown). During the growing season, the assimilation of CO₂ by the plants (A_n term in equation 1.12) drives the NEE. Differences in the A_n calculation together with differences in the meteorological forcing used explain quite high discrepancy between the four data sets. Nevertheless, the marked seasonality associated to the African monsoon is well reproduced by the three models. Concerning the wet season NEE (table 2), CTESSEL open-loop and CASA are the lower, SIB shows the higher value whereas CTESSEL analysis is in between. In addition, compared to the climatologies SIB and CASA, CTESSEL is able to simulate an interannual variability. Over the period of study, the timing of the growing season differs: SIB/CASA and CTESSEL are in phase for 2001 but are strongly shifted in 2004. The amplitude of the NEE seasonal signal shows also a strong interannual variability. Over savannah, SIB and CASA are in good agreement whereas CTESSEL (analysis and open-loop) has a stronger seasonal signal with higher respiration rates during the dry season and higher CO₂

absorption during the growing season. Compared to SIB and CASA, the CTESSEL seasonal cycle is also characterized by a CO₂ assimilation peak either in phase (in 2001 and 2002) or delayed. This delay reaches two months in 2003. The decrease of the CO₂ assimilation after the peak occurs systematically later for CTESSEL than for SIB and CASA. These shifted seasonal cycles are in accordance with the delayed LAI cycle already identified above. The LAI data assimilation partly improves this shifted CO₂ fluxes. Over Forest, SIB and CASA are slightly shifted in time but they both show a strong biennial signal (with CO₂ absorption during spring and autumn and CO₂ drop during summer and winter) that is not reproduced by the CTESSEL model.

As a conclusion, CTESSEL exhibits CO₂ fluxes in relatively good agreement with two data sets extensively used as boundary condition for CO₂ transport model. After data assimilation, CTESSEL NEE is even in better timing with SIB and CASA than CTESSEL open-loop. At ECMWF, CTESSEL is intended to replace the CASA climatology used for the CO₂ surface fluxes. Its ability to simulate interannual variability of surface carbon fluxes is expected to bring a significant improvement on the actual system.

8. Conclusions and perspectives

The ECMWF land surface model has been recently modified to include an interactive vegetation module able to simulate CO₂ Net Ecosystem Exchanges and to diagnose Leaf Area Index (LAI). It is named CTESSEL for Carbon-TESSSEL. A data assimilation system to analyse LAI from satellite products into this model is presented and evaluated over West Africa from 2001 to 2005. Prior to the data assimilation, two satellite derived LAI data sets are compared to the CTESSEL open-loop simulations over a large window covering West Africa: the CYCLOPES and the MODIS LAI. The satellite products are significantly lower than the CTESSEL simulations over the whole study region. The CTESSEL simulations exhibit a shifted phenology with regards to satellite derived LAI. The growth is quite low and the senescence is late over Sahel and savannah. The negative bias of the satellite derived LAI is reduced thanks to a histogram adjustment method. Finally, the MODIS LAI are chosen as input for the data assimilation system as they are available over the entire period of study and they fit better to the CTESSEL simulations after debiasing. The data assimilation method is a simplified 2DVAR. It presents the advantage to avoid the development of the adjoint and the tangent linear of the CTESSEL model.

This approach is firstly evaluated locally on the sahelian site of the AMMA project located in Mali. With regards to the open-loop, the CTESSEL simulations constrained by the analysis of LAI are more in agreement with the ground measurements in terms of LAI and above ground biomass. In particular, the delayed growth of the open-loop is corrected thanks to the data assimilation system. In terms of fluxes, open-loop and analysis are close until the LAI peak. Afterwards, the open-loop maintains a high evapotranspiration rate and NEE thanks to a high LAI. In contrast, the analysis has reduced the LAI to zero: the evapotranspiration and NEE are much lower and close to null. Additionally, the order of magnitude of NEE and LE (with or without data assimilation) are in good agreement with measurements performed by other teams over similar ecosystems. Nevertheless, the soil respiration parameterization of the CTESSEL model is too rough to capture the dynamics of respiration of the sahelian soils and must be improved in a near future.

The simplified 2DVAR approach is applied to West Africa in a second part. It is shown that the data assimilation performs well in terms of LAI with analysis closer to the satellite derived LAI than the forecast. Therefore, the slow growth and the late senescence of the CTESSEL open-loop LAI simulations highlighted

above are strongly improved thanks to the data assimilation system. In terms of fluxes, the impact of the data assimilation is much higher on CO₂ fluxes than on LE. Likewise, using two different meteorological forcing, it is demonstrated that the seasonal and intra-seasonal dynamic of the LE fluxes are mainly influenced by the rainfall events (distribution and quantity) whereas the main driver of NEE is LAI. As a conclusion, whereas an interactive vegetation module seems a requisite to simulate CO₂ fluxes variability, being able to simulate a LAI responding to climate is of secondary importance for LE.

At ECMWF, the CTESSEL model is intended to provide the surface fluxes of natural CO₂ instead of the CASA data sets used at present within the GEMS project and will participate to the TRANSCOM experiment. The CTESSEL NEE comparison to the SIB 3.0 and CASA data sets shows that CTESSEL gives NEE values of the same order of magnitude as the two data sets but that the simulated seasonal evolution is slightly shifted. The CTESSEL LAI are currently compared to other land surface models within the ALMIP project. Finally, within the frame of the GEOLAND 2 program to begin in 2008, the soil parameterization will be improved in a near future and the simplified 2DVAR for the analysis of LAI will be evaluated at the global scale.

9. Acknowledgement

“This study was co-funded by the European Commission within FP6, in the framework of the GEOLAND integrated GMES project on land cover and vegetation.” We wish to acknowledge Jim Randerson and colleagues, and Ian Baker, Scott Denning and colleagues for making available a free on-line version of global NEE fluxes of CO₂ from CASA and SiB.3 process models, respectively.

The EU-funded AMMA project is acknowledged for providing the forcing and the special observations for the validation and assimilation. Many scientific feedbacks have been provided in particular from AMMA-ALMIP initiative conducted at CNRM and CESBIO for which we wish to thank in particular Aaron Boone and Patricia Derosnay.

10. References

- ALMIP, 2006, the AMMA Land surface Model Intercomparison Project, http://www.cnrm.meteo.fr/amma-moana/amma_surf/almip/index.html.
- Arora, V.K., 2002, Modelling vegetation as a dynamic component in soil-vegetation-atmosphere-transfer schemes and hydrological models, *Reviews of Geophysics*, **40**(2),1006, 10.1029/2001RG000103.
- Balsamo G., Bouyssel F., and Noilhan J., 2004, A simplified bi-dimensional variational analysis of soil moisture from screen-level observations in a mesoscale numerical weather-prediction model. *Q. J. R. Met. Soc.*, **130A**, 895-915.
- Baret F., Hagolle O., Geiger B., Bicheron P., Miras B., Huc M., Berthelot B., Niño F., Weiss M., Samain O., Roujean J.L. and Leroy M., 2007, LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm, *Remote Sens. Environ.*, in press.
- Beljaars, A.C.M., and P. Viterbo, 1999, Soil moisture-precipitation interaction: Experience with two land surface schemes in the ECMWF model. *Global energy and water cycles* (ed. by K. Browning and R. Gurney), Cambridge University Press, Cambridge, 223-233.

- Calvet J.C., Noilhan J., Roujean J.-L., Bessemoulin P., Cabelguenne M., Olioso A., Wigneron J.-P., 1998, An interactive vegetation SVAT model tested against data from six contrasting sites. *Agric. For. Meteorol.*, **92**, 73-95.
- Calvet, J.-C., 2000, Investigating soil and atmospheric plant water stress using physiological and micrometeorological data, *Agric. For Meteorol.*, **103**, 229-247.
- Calvet J.C., and Soussana, J.F., 2001, Modelling CO2 enrichment effects using an interactive vegetation SVAT scheme. *Agric. For. Meteorol.*, **108**, 129-152.
- Denning, A. S., G. J. Collatz, C. Zhang, D. A. Randall, J. A. Berry, P.J. Sellers, G. D. Colello, and D. A. Dazlich, 1996, Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model. Part 1: Surface carbon fluxes. *Tellus*, **48B**, 543 - 567.
- Duchemin B., Berthelot B., Dedieu G., Leroy M., Maisongrande P., 2002, Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT. II - Validation of an operational method on actual data sets, *Remote Sensing of Environment*. **81**, no. 1, pp. 101-113.
- Falk U., Brümmer C., Brüggemann N., Wassmann R. and Szarzynski J., 2007, Fluxes of Carbon and energy above a natural savannah in Burkina Faso, West Africa, *Geophysical Research Abstracts*, **9**, 09302.
- GEMS, 2005, Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data (http://www.ecmwf.int/research/EU_projects/GEMS/).
- GEOLAND, 2004, Integrated GMES project on Land Cover and Vegetation, <http://www.gmes-geoland.info>.
- Gibelin, A., J. Calvet, J. Roujean, L. Jarlan, and S. O. Los, 2006, Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products, *J. Geophys. Res.*, **111**, D18102, doi:10.1029/2005JD006691.
- Gu Y., Belair S., Mahfouf J.F. and Deblonde G., 2006, Optimal interpolation analysis of leaf area index using MODIS data, *Remote Sensing of Environment*, doi:10.1016/j.rse.2006.04.021.
- Hanan N.P., Kabat P., Dolman A.J. and Elbers J.A., 1998, Photosynthesis and carbon balance of a Sahelian fallow savanna, *Global change biology*, **4**, 523-538.
- Hiernaux, P., 1984, Distribution des pluies et production herbacée au Sahel: une méthode empirique pour caractériser la distribution des précipitations journalières et ses effets sur la production herbacée. Premiers résultats acquis dans le Sahel malien. *Document du programme Zones arides et semi-arides, CIPEA Bamako et ILCA Addis Ab eba*. 594.
- Jacobs, C. M. J., 1994, *Direct impact of atmospheric CO2 enrichment on regional transpiration*. PhD thesis, Wageningen Agricultural University.
- Kabat P., Dolman A.J., Elbers J.A., 1997, Evaporation, sensible heat and canopy conductance of fallow savannah and patterned woodland in the Sahel, *J. of Hydrology*, **188/189**, 494-515.
- Kimes D.S., Knyazikhin Y., Privette J.L., Abuelgasim A.A. and Gao F., 2000, Inversion methods for physically based models, *Remote sensing reviews*, **18**, 381-439.
- Kucukkaraca. E. and M. Fisher, 2006, Use of analysis ensembles in estimating flow-dependent background error variances, ECMWF technical memoranda n°492 (<http://www.ecmwf.int>), 16 pages.

- Lacaze R., 2004, Algorithm theoretical Basis Document (ATDB), customization for LAI, fAPAR, fcover and albedo, http://postel.mediasfrance.org/IMG/pdf/CYCL_ATBD-DirectionalNormalisation_I2.0.pdf.
- Lafont S., Voogt M., Beljaars A., Van den Hurk B.J.J.M., Jarlan L., Balsamo G., 2007, Description and evaluation of the Carbon version of TESSEL: CTESSSEL model, ECMWF technical memorandum, 37 p.
- Le Hou  rou H.N., 1989, The Grazing Land Ecosystems of the African Sahel, *Ecological Studies* 75, Springer-Verlag, Berlin, 282 pp.
- Lo Seen D., Mougin E., Rambal S., Gaston A., Hiernaux P., 1995, A regional Sahelian grassland model to be coupled with multispectral satellite data. II. Towards its control by remotely sensed indices., *Remote Sens. Environ.*, **52**, 194-206.
- Masson, V., J.L. Champeaux, F. Chauvin, C. M  riguet and R. Lacaze, 2003, A global database of land surface parameters at 1km resolution for use in meteorological and climate models. *J. Climate*, **16**, pp 1261-1282.
- Mu  oz Sabater J., Jarlan L., Calvet J.C., Bouyssel F., De Rosnay P., 2007a, From near surface to root zone soil moisture using different assimilation techniques, *Journal of hydrometeorology*, **8**, 194-206.
- Mu  oz Sabater J., R  diger C., Calvet J.C., Fritz N., Jarlan L. and Kerr Y., 2007b, Joint assimilation of surface soil moisture and LAI observations using a simplified 1D-VAR: the smosrex case study, *Agricultural and Forest Meteorology*, in revision.
- Myneni R. B., Nemani R.R., Running S.W., 1997, Estimation of global leaf area index and absorbed par using radiative transfer models, *IEEE Trans. Geosc. Rem. Sensing*, **35**, 1380 - 1393, doi: 10.1109/36.649788
- Philippon, N., and B. Fontaine, 2001, The relationship between the Sahelian and previous second Guinean rainy seasons: a monsoon regulation by soil wetness, *Annales Geophysicae*, **20**, 4, 575-582.
- Prince S.D., 1991, Satellite remote sensing of primary production : comparison of results for Sahelian grasslands 1981-1988, *Int. J. Remote Sensing*, **12**, n   6, p. 1301-1311.
- Randerson, J.T., M.V. Thompson, T.J. Conway, I.Y. Fung, and C.B. Field, 1997, The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochemical Cycles*, **11**: 535-560.
- Roujean, J.-L., and R. Lacaze, 2002, Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation, *Journal of Geophysical Research*, **107**(D12), 4150, doi:10.1029/2001JD000751.
- Shukla, J., and Mintz, Y., 1982, Influence of Land-Surface Evapotranspiration on the Earth's Climate. *Science*, **215**, 1498-1501.
- Sivakumar M.V.K., 1990, Exploiting rainy season potential from the onset of rains in the Sahelian zone of West Africa, *Agricultural and Forest Meteorology*, **51**, pp. 321-332.
- Tian, Y., Zhang, Y., Knyazikhin, Y., Myneni, R. B. & Running, S. W., 2000, Prototyping of MODIS LAI/FPAR algorithm with LASUR and Landsat data. *IEEE Transaction on Geoscience and Remote Sensing*, **38**(5), 2387-2401.

- Tucker, C.J., J. E. Pinzon, M. E. Brown, D. Slayback, E. W. Pak, R. Mahoney, E. Vermote and N. El Saleous, 2005, An Extended AVHRR 8-km NDVI Data Set Compatible with MODIS and SPOT Vegetation NDVI Data. *International Journal of Remote Sensing*, **26**:20, pp 4485-5598.
- Tucker C.J., Fung Y.I., Keeling C.D. and Gammon R.H., 1986, Relationship between atmospheric CO₂ variations and a satellite derived vegetation index, *Nature*, **319**, 195-199.
- Tucker C.J., Vanpraet C.L., Sharman M. J. and Van Ittersum G., 1985, Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980-1984, *Remote Sensing of Environment*, **17**, pp. 223-249.
- Van den Hurk, B. J. J. M., P. Viterbo, A. C. M. Beljaars, and A. K. Betts, 2000, Offline validation of the ERA40 surface scheme. ECMWF Tech. Memo 295, ECMWF, 43 pp.
- Van den Hurk B.J.J.M., Viterbo P. and Los S.O., 2003, Impact of Leaf area index seasonality on the annual land surface evaporation in a global circulation model, *Journal of Geophysical Research*, **108**, N0 D6, 4191, doi:10.1029/2002JD002846.
- Veenendaal E. M., Kole O., Lloyd J., 2004, Seasonal variation in energy fluxes and carbon dioxide exchange for a broad-leaved semi-arid savanna (Mopane woodland) in Southern Africa, *Global Change Biology*, **10**, 318-328, doi:10.1046/j.1529-8817.2003.00699.x.
- Verhoef A., Allen S.J., De Bruin H.A.R., Jacobs C.M.J. and Heusinkveld B.G., 1996, Fluxes of carbon dioxide and water vapour from a Sahelian Savanna, *Agricultural and Forest Meteorology*, **80**, 231-248.
- Viovy N., O. Arino, A.S. Belward, 1992 The Best Index Slope Extraction (BISE) – a method for reducing noise in NDVI time series, *International Journal of Remote Sensing* **13** (8): 1585-1590.
- Voogt, M.H., B.J.J.M. van den Hurk and C.M.J. Jacobs, 2006, *The ECMWF land surface scheme extended with a photosynthesis and LAI module tested for a coniferous forest site* Agric. and Forest Met..
- Wallace J.S. and Holwill C.J., 1997, Soil evaporation from Tiger-bush in south-west Niger, *Journal of Hydrology*, **189**, pp. 426-442.
- Weiss M., Baret F., Smith G. J., Jonckheere I., and Coppin P., 2004, Review of methods for in situ leaf area index (LAI) determination: Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*, **121**(1-2):37--53, 2004.
- Weiss M., Baret F., Garrigues S., Lacaze R., 2007, LAI dans fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products, *Remote Sens. Environ.*, in press.
- Zeng, N., J. D. Neelin, K.-M. Lau, and C. J. Tucker, 1999, Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science*, **286**, 1537-1540.

Appendix 1: the CTESSEL model

The photosynthesis and stomatal resistance model

The canopy resistance is calculated from photosynthesis, which is the: net CO₂ assimilation (A_n) by the canopy. A_n is calculated as a function of different environmental factors. First, CO₂ assimilation limited by the air CO₂ concentration is determined via a saturation equation:

$$A_m = A_{m,\max} \left(1 - e^{(-g_m (C_i - \Gamma) / A_{m,\max})} \right) \quad (1.1)$$

where $A_{m,\max}$ is the maximum net CO₂ assimilation, g_m is the mesophyll conductance, C_i is the CO₂ concentration in the leaf and is the CO₂ concentration at which assimilation compensates respiration, called CO₂ compensation concentration. $A_{m,\max}$ depends on temperature via a Q₁₀ function. The internal CO₂ concentration C_i , is directly derived from the CO₂ concentration in the air C_s . It is controlled by the air humidity via the specific humidity deficit of the air D_s . If the deficit exceeds the maximum deficit tolerated by the vegetation D_{\max} , the plant closes its stomata. D_{\max} is vegetation dependent. The CO₂ assimilation limited by CO₂ concentration is further limited by radiation by:

$$A_n = (A_m + R_d) \left[1 - e^{\frac{-\varepsilon I_a}{(A_m + R_d)}} \right] - R_d \quad (1.2)$$

where I_a is the photosynthetic active radiation (PAR), ε is the initial quantum use efficiency and R_d is the dark respiration. R_d is parameterized simply as a function of A_m .

The stomatal conductance to CO₂, g_{sc} , is estimated using a flux-gradient relationship, modified to account for the effect of a specific humidity deficit on stomatal aperture. The first guess g_{sc}^* is given by:

$$g_{sc}^* = \frac{A_n - A_{\min} \left(\frac{D_s}{D_{\max}} \frac{A_n + R_d}{A_m + R_d} \right) + R_d \left(1 - \frac{A_n + R_d}{A_m + R_d} \right)}{C_s - C_i} \quad (1.3)$$

where A_{\min} represents the residual photosynthesis rate (at full light intensity) associated with cuticular transfers when the stomata are closed because of a high specific humidity deficit. The diffusion of CO₂ interacts with that of water vapour. The first guess of the stomatal conductance to CO₂ is corrected for this interaction and calculated interactively to refine the estimation. Finally, the stomatal conductance to water vapour g_s is given by:

$$g_s = 1.6 g_{sc} + g_c \quad (1.4)$$

where g_c is the cuticular conductance (a vegetation dependent parameter).

The soil stress parameterization is based on a meta-analysis of several herbaceous and woody vegetation types (Calvet, 2000). The meta-analysis shows relationships between g_m and D_{\max} for low vegetation and between g_m and f_0 for high vegetation. Furthermore, it seems that plants react in two different ways to soil moisture stress. There are plants that try to avoid stress, by reducing the evaporation via stomatal regulation. This stress strategy is typified as defensive. Others apply an offensive strategy in order to resist stress, by a

more efficient root water-uptake or a more rapid growing cycle. Among species within the 7 vegetation classes of CTESSEL often both strategies occur. Therefore, it is not easy to generalize the strategy for each class. It seems most likely that coniferous forest has a defensive strategy, whereas an offensive strategy is assigned to the other classes. Further details on the soil stress parameterization can be found in Voogt et al. (2007). The net CO₂ assimilation calculated at the leaf scale is upscaled at the canopy scale assuming that leaf parameters do not vary within the canopy. The attenuation of the incoming shortwave radiation in the canopy is computed thanks to a simple radiative transfer model (Calvet et al., 1998).

10.1.1. The vegetation growth model

The interactive LAI is based on biomass evolution due to photosynthetic activity. The biomass module simulates growth and mortality of the vegetation. The growth of active biomass B is based on the accumulated net CO₂ assimilation over the previous day $A_{n,day}$ and the LAI is obtained from the biomass following:

$$\alpha_B = \frac{B}{LAI} \quad (1.5)$$

In reality, α_B depends on climate (temperature and CO₂ concentration) and nitrogen fertilisation. In order to account for plant morphology, the nitrogen dilution concept is applied for the biomass evolution. The plant N decline model is a well-established agronomical law relating the plant N in non-limiting N-supply conditions to the accumulated above-ground dry matter. The critical plant N is the value of N maximizing growth, and this value decreases for increasing biomass accumulation following a negative power law. The basis of the model is that the metabolic component B of the plant biomass is related to total biomass B_T through an allometric logarithmic law (Calvet and Soussana, 2001). In CTESSEL, the metabolic biomass component is identified as the active biomass and The relationship between active biomass B and total aboveground biomass B_T is:

$$B_T = \left(\frac{B}{c}\right)^{\frac{1}{1-a}} \quad (1.6)$$

where a and c are constant parameters. The total aboveground biomass consists of the active biomass reservoir and the structural aboveground reservoir B_s , which can be considered as the "living" structural biomass, like the stem. For forests, wood is a dead reservoir and does not contribute to B_s . Within the nitrogen dilution model a relationship between the leaf area ratio LAR and the aboveground nitrogen concentration N_T is applied:

$$LAR = \frac{LAI}{B_T} = eN_T + f \quad (1.7)$$

where e and f are called plasticity parameters and are derived per vegetation type. Eq.1.7 can be used as a closure equation to estimate α_B :

$$\alpha_B = \frac{1}{eN_a + \frac{f}{cB_T^{-a}}} \quad (1.8)$$

where N_a is the nitrogen concentration in the active biomass. It depends on vegetation type and on the nitrogen fertilisation. For further details and derivations see Calvet and Soussana (2001). In this way, α_b has

become a model variable depending on B_T . However, for global simulations, it is desirable to keep α_b as a constant parameter in order to let α_B represent rather intrinsic plant characteristics denoting a biological adaptation to average climate and growing conditions (Calvet and Soussana, 2001). For that purpose, Eq.1.8 can only be solved by iteration. Moreover, LAR and N_T data to derive the plasticity parameters by regression is lacking. However, data is available for leaves in the form of the specific leaf area SLA and the nitrogen content in leaves N_L :

$$SLA = \frac{LAI}{B_L} = eN_L + f \quad (1.9)$$

Both the iteration issue and the availability of data to derive e and f give rise to modify the nitrogen dilution module. Eq.1.8 is simplified by considering α_B as the ratio of the biomass of green leaves to LAI:

$$\alpha_B = \frac{1}{SLA} = \frac{1}{eN_L + f} \quad (1.10)$$

10.1.2. Soil respiration and NEE parameterization

Since the biomass model is not coupled to a soil model, soil respiration needs to be parameterized in another way. In CTESSEL the respiration is calibrated in order to simulate a zero net CO_2 exchange over a multi-year period. The CO_2 ecosystem respiration is split into two terms. The first is the dark respiration R_d (parameterized from A_m). The second respiration term represents both heterotrophic respiration from the soil and autotrophic respiration from the above- and below ground structural biomass and named, hereafter, $R_{soilstr}$. A Q_{10} function is used for its parameterization:

$$R_{soilstr} = R_{eco} - R_d = R_0 Q_{10}^{((T_{soil} - 25)/10)} \quad (1.11)$$

where R_0 is the reference respiration at 25°C, T_{soil} is the temperature of the 2nd soil layer and Q_{10} is fixed at 2.0. R_0 is determined per vegetation type in each grid box, assuming equilibrium between multi-annual net CO_2 assimilation, harvest and residual respiration.

Finally, the net ecosystem CO_2 exchange (NEE) is given by:

$$NEE = A_n - R_{soilstr} \quad (1.12)$$