



**HAL**  
open science

## Using coarse remote sensing radar observations to control the trajectory of a simple Sahelian land surface model

Lionel Jarlan, Éric Mougin, P. Mazzega, Marc Schoenauer, Yann Tracol,  
Pierre Hiernaux

► **To cite this version:**

Lionel Jarlan, Éric Mougin, P. Mazzega, Marc Schoenauer, Yann Tracol, et al.. Using coarse remote sensing radar observations to control the trajectory of a simple Sahelian land surface model. *Remote Sensing of Environment*, 2005, 94 (2), pp.269-285. 10.1016/j.rse.2004.10.005 . ird-00391596

**HAL Id: ird-00391596**

**<https://ird.hal.science/ird-00391596>**

Submitted on 24 Jul 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Using coarse remote sensing radar observations to control the trajectory of a simple Sahelian land surface model

L. Jarlan<sup>a,\*</sup>, E. Mougin<sup>a</sup>, P. Mazzega<sup>b</sup>, M. Schoenauer<sup>c</sup>, Y. Tracol<sup>a</sup>, P. Hiernaux<sup>d</sup>

<sup>a</sup>Centre d'Etudes Spatiales de la Biosphère, 18 avenue Edouard Belin, 31401 Toulouse Cedex 4, France

<sup>b</sup>Laboratoire d'Etudes en Géophysique et Océanographie Spatiale, 18 avenue Edouard Belin, 31401 Toulouse Cedex 4, France

<sup>c</sup>Projet FRACTALES, I.N.R.I.A. Rocquencourt, B.P. 105, 78153 LE CHESNAY Cedex, France

<sup>d</sup>Tropenzentrum Universität Hohenheim (790), 70593 Stuttgart, Germany

---

## Abstract

In the Sahel, land surface processes are significantly interacting with climate dynamics. In this paper, we present an original method to control a simple Sahelian land surface model coupled to a radiative transfer model (RTM) on the basis of ERS wind scatterometer (WSC) observations. In a first step, a sensitivity study is implemented to identify those parameters of the land surface model that can be estimated through the assimilation of WSC data. The assimilation scheme relies on evolution strategies (ES) algorithm that aims at solving the parameter evaluation problem. These algorithms are particularly well suited for complex (nonlinear) inverse problems. The assimilation scheme is applied to several study sites located in the Sahelian mesoscale site of the African Monsoon Multidisciplinary Analysis Project (Gourma region, Mali). The results are compared with ground observations of herbaceous mass. After the WSC data assimilation, the simulated herbaceous mass curves compare well with observations [187 kilogram of dry matter per hectare (kg DM/ha) of average error]. The simulated water fluxes exhibit a behaviour in agreement with ground measurements performed over similar ecosystems during the Hapex Sahel experiment. The accuracy of estimated herbaceous mass and water fluxes resulting from uncertainties on climatic forcing variable is evaluated using a stochastic approach. The average error on the herbaceous mass values mainly depends on the rainfall estimate accuracy and ranges from 139 to 268 kg DM/ha that compares well with a previous study based on the sole inversion of the radiative transfer model. Finally, this study underlines the need for a multispectral assimilation approach to get a better constraint on water fluxes estimation.

© 2004 Elsevier Inc. All rights reserved.

*Keywords:* Wind scatterometer; Sahel; Data assimilation; Radiative transfer modelling; Vegetation functioning model; Parameter identification; Evolution strategies

---

## 1. Introduction

The land surface component of the hydrological cycle has a major role on the climate of the earth through the exchanges of water and energy at the soil–vegetation–atmosphere interface (Shukla & Mintz, 1982). Over the Sahel, characterized by a high water recycling rate (Brubaker et al., 1993), the water fluxes in particular are of prime importance to the understanding and forecasting of the climate (Nicholson,

2000). Vegetation cover modifies the partition of latent fluxes at the surface between soil evaporation and plant transpiration and thus alters the water content of the atmospheric boundary layer available for the development of the convective rainfall events (Mahrt, 2000). Within this context, Taylor and Lebel (1998) have found observational evidence of the direct impact of this partition on the local persistence of convective events in the boundary layer associated to rainfall in this region. In particular, vegetation transpiration, by delaying the return of water to the atmosphere, favours the upkeep of convective movements between two rainfall events.

Regarding the role of land surface processes on climate, increasing realistic land surface models have been incorpo-

---

\* Corresponding author. Tel.: +33 561 558537; fax: +33 561 558500.

E-mail address: lionel.jarlan@cesbio.cnrs.fr (L. Jarlan).

rated into general circulation models (GCM) by climate modelling communities since Manabe's pioneering work in 1969. A land surface model forced by meteorological data and surface-specific characteristics (vegetation types and soil properties) simulates processes of the surface functioning such as vegetation dynamics, water, energy and carbon exchanges, soil water dynamics, etc. (e.g., Dickinson, 1995; Polcher et al., 1995). The ability of such models to simulate realistic variations of surface properties as well as exchanges at the soil–vegetation–atmosphere interface faces several limitations: (a) forcing variables such as rainfall or available energy exhibits strong spatiotemporal variability; (b) no suitable representation is available for the key land surfaces processes at large scale; (c) for those processes adequately parameterized, there is no efficient and generic methods to aggregate the available information at a large scale (Kabat, Hutjes et al., 1997; Sellers et al., 1997). As a consequence, available land surface models have considerable difficulties to correctly simulate surface water and energy partitioning and related processes, which limits climate prediction. This has motivated several recent studies that aim at correcting certain state variables of the land surface models via the assimilation of remote sensing data (Boulet et al., 2002; Cayrol et al., 2000; Crow & Wood, 2003; Reichle et al., 2001, 2002).

Today, limited to the optical domain, coarse resolution remote sensing sensors enable surface characteristics such as vegetation cover or leaf area index (LAI) to be monitored over large areas or even globally with a high temporal repetitivity. In addition, radar data acquired by ERS wind scatterometer (WSC) instruments on board ERS-1 and ERS-2 satellites have shown strong potentials for monitoring the Sahelian environment, mainly vegetation living mass and soil moisture (Frison & Mougouin, 1996a; Frison et al., 1998; Jarlan, Mougouin et al., 2002; Magagi & Kerr, 1997; Wagner & Scipal, 2000; Woodhouse & Hoekman, 2000). The sensitivity of scatterometers originates from the drastic change of dielectric properties of the surface associated with vegetation development and soil moistening that strikes the Sahelian belt between dry and wet season. However, this dataset has never been used within an assimilation procedure.

This study demonstrates the potentialities of coarse radar observations to simultaneously control the simulated vegetation dynamics and water budget of a simple land surface model restricted to Sahelian grasslands. The paper is organised as follows. The second section provides a brief description of the data and models used. The assimilation methodology is detailed in the third section. The fourth section deals with assimilation effects on ecosystem functioning simulations (vegetation production and water fluxes). Comparison with available ground measurements performed over many growing seasons and error bar estimates is also presented. Finally, conclusions and perspectives are drawn. This work is part of the African Monsoon Multidisciplinary Analysis project dedicated to a

better understanding of the interannual and seasonal dynamics of the African Monsoon (AMMA, 2001).

## 2. Data and models

### 2.1. The wind scatterometer data set

The ERS wind scatterometer (WSC) instrument provides backscattering coefficient  $\sigma^0$  measurements at 5.3 GHz and VV polarisation. Data were collected from August 1991 to January 1995 onboard ERS-1 and from 1995 to December 2000 onboard ERS-2. The spatial resolution is about 50 km, with incidence angles comprising between 18° and 59°. Data are provided with a 0.25°×0.25° pixel size. A more detailed description of the instrument capabilities can be found in Frison and Mougouin (1996b). In the present study, to avoid data discrepancy, only observations acquired between 40° and 55° are kept. They are linearly normalized at a 45° incidence angle. The average number of available data is about 1.8 data/decade. The standard deviation (S.D.) of data error is estimated to be 0.55 dB over the Sahel at a 45° incidence angle (Jarlan, Mazzega et al., 2002). For lack of precise information, data are assumed to be contaminated by a white noise.

### 2.2. The land surface model

The Sahelian Transpiration Evaporation and Production (STEP) model is a simple land surface model for Sahelian regions that was originally designed for use with remote sensing data (Lo Seen et al., 1995; Mougouin et al., 1995). The STEP model describes the main relevant processes associated with herbage functioning and soil water dynamics in a Sahelian environment. Processes of the soil–plant–atmosphere system, such as water fluxes, photosynthesis, respiration, plant growth and senescence are simulated with a daily time step. The model is driven by standard daily meteorological variables (rainfall, incident global radiation, wind speed, minimum and maximum air temperature and mean relative air humidity), among which the most important are rainfall and global radiation. STEP includes two submodels dealing with plant growth and water fluxes. The present growth submodel simulates processes of plant emergence, photosynthesis, carbon allocation using two compartments (aerial vegetation and roots), plant growth and senescence. Plant transpiration and soil evaporation are simulated by the water dynamics submodel using the Penman–Monteith equations (Monteith, 1965). Water interception by vegetation canopy is neglected regarding the low vegetation cover fraction (Vcf) in this region. The soil water dynamics is described using a simple four-layer water bucket model (Manabé, 1969). STEP simulates on a daily basis the structural variables of the vegetation canopy [canopy height  $h_c$ , total herbaceous mass  $B_m$ , leaf area index (LAI) and vegetation cover fraction (Vcf)], the

vegetation water content  $H_p$ , the soil water content  $H_v$  and water fluxes. The main equations of the STEP model are given in Appendix A, and a further description can be found in Mougín et al. (1995).

### 2.3. The radiative transfer model

The STEP model has already been coupled with a radiative transfer model (RTM) in the radar domain (Frison et al., 1998; Jarlan, Mougín et al., 2002). This model simulates the interaction between the electromagnetic wave emitted by the scatterometer and the surface components, namely, soil and vegetation. The herbage canopy is modelled as a collection of discrete randomly oriented scatterers above a rough surface. For the calculations, the herb components are modelled as elliptic discs. Tree cover is neglected (Jarlan, Mougín et al., 2002). The RTM links the backscattering coefficient to:

- the water content of the upper soil profile  $H_v$  through the soil dielectric constant,
- the vegetation water content  $H_p$  through the vegetation dielectric constant,
- the standing herbaceous mass  $B_m$  and canopy height  $h_c$  through the scatterer density  $n_0$ ,
- the scatterer geometrical parameters (dimension and orientation).
- the surface roughness.

The scatterer density  $n_0$  is derived from the vegetation volume fraction  $f_v$  and the volume of an elementary scatterer  $V_{\text{scat}}$  as (Frison et al., 1998):

$$n_0 = \frac{f_v}{V_{\text{scat}}} \quad (1)$$

$$\text{with } f_v = 0.0001 B_m \frac{[\psi + H_p(1 - \psi)]}{h_c \rho_e (1 - H_p)} \quad (2)$$

where  $\psi$  is the ratio between the water density  $\rho_e$  and the dry matter density. Table 1 lists the RTM input parameters (the main equations of the RTM model can be found in Frison et al., 1998).

For lack of a dense ground station network to describe rainfall spatial variability, the simulated soil water content of the upper soil profile from the STEP model is averaged over

a 15-day period to get a representative value at the 50-km scale following Frison et al. (1998). Previous studies show that the combined RTM–STEP models enable interannual ERS measurements to be well simulated. Details about the coupled models can be found in Frison et al. (1998) and Jarlan, Mougín et al. (2002).

### 2.4. Study region and ground observations

The study region is located in the Gourma region in Mali. The Malian Gourma belongs to the Sahelian zone and extends to the south of the Niger River between Timbuktu and Gao down to the border with Burkina-Faso (14°N/17°N and 2°W/1°W). This is mainly a pastoral region bracketed by the 500 and 150 mm/year isohyets. Rainfall distribution is strictly monomodal starting in June and ending in September, with a maximum in August. Over the year, mean daily global radiation is about 18 MJ m<sup>-2</sup> d<sup>-1</sup>. The mean annual temperature is about 29 °C, with maximum and minimum monthly values of 41 and 15 °C occurring in April and January, respectively.

Crops are almost absent from the Gourma region apart from the southern border. Natural vegetation is composed of a herbage layer and a sparse tree layer. The herbage layer is dominated by annual plants among which C4 graminoids prevail. For the considered sites, soils are largely sandy (85–95%), poor in organic matter content (0.1–0.2%) and acidic (pH 5–6). The development of herbage plants starts after the first rains in June or July, and, unless the plants wilt before maturity by lack of rain, the senescence follows the fructification which matches with the end of the rainy season.

Herbaceous mass measurements are collected between one and three times over the growing season over four 25\*25 km<sup>2</sup> sites and are expressed in kilogram of dry matter per hectare (kg DM/ha). These sites, located along a north–south transect, cover the whole bioclimatic gradient of the Sahelian zone, as defined by Le Houérou (1989). Over the study period, the herbaceous mass was measured during the 1992, 1993, 1999 and 2000 seasons (21 points). The uncertainty of herbaceous mass measurement at that scale is estimated about 30% (Jarlan et al., 2003). Location of the study sites and herbaceous mass measurements are given in Table 2. Further description of the study sites and field measurements is available in Hiernaux and Justice (1986) and Tracol et al. (submitted for publication).

Table 1  
List of the input variables of the radiative backscattering model

Symbol	Definition	Units	Value	Source
$x, y, z$	Scatterer dimension	cm		Measurements on typical C4 grass (Frison et al., 1998)
$\alpha, \beta, \gamma$	Scatterer orientation	radian	[0;2 $\pi$ ]	Measurements on typical C4 grass (Frison et al., 1998)
$n_0$	Scatterer density	m <sup>-3</sup>	–	Function of herbaceous mass, canopy height and vegetation water content
$\epsilon_v$	Dielectric constant of scatterers	–	–	Function of vegetation water content (Ulaby & El-Rayes, 1987)
$\epsilon_s$	Dielectric constant of soil	–	–	Function of soil water content (Hallikainen et al., 1985)
$s$	Soil roughness	cm	[0.001;0.05]	Fitted using backscattering coefficient acquired during dry season (Frison et al., 1998)

Table 2

Location of the study sites and herbage mass measurements (the measurement date is given as day of the year -DOY-)

	Location		Season 1992		Season 1993		Season 1999		Season 2000	
	Latitude	Longitude	Date (DOY)	$B_m$ (kg DM/ha)						
Rharous	16.5°N	1.5°W	267	375	254	505	270 219	1550 60	246	650
Gossi	16.0°N	1.4°W	267 215	329 163	253	250	261 213	1630 99	244	200
Hombori	15.3°N	1.5°W	234 263	543 1061	251	744	264	1380	213 250	190 990
Seno	14.6°N	2.9°W	264	1434	249	2135	271	1634	289	1240

Rainfall data are recorded by four meteorological stations located near each study site: Gourma Rharous (16.9°N, 1.9°W), Gossi (15.8°N, 1.3°W), Hombori (15.3°N, 1.7°W) and Diankabou (14.6°N, 3.1°W). The other forcing climatic variables for the STEP model, namely, global radiation, temperature, air humidity and wind speed, are extracted from the climatological atlas of Morel (1992).

### 3. Assimilation methodology

Regarding the numerous parameters of the coupled RTM–STEP models, a sensitivity study is first conducted to identify those parameters that can be constrained via the WSC data assimilation. The assimilation methodology is then described.

#### 3.1. Sensitivity study

The coupled RTM–STEP models have numerous parameters describing the surface functioning and the physical properties of the observed surface. From previous studies, parameters of the RTM (see Table 1) are set to values defined in Frison et al. (1998) and Jarlan, Mougin et al. (2002). Scatterer dimension and orientation parameters have values of typical Sahelian annual herbage species. Soil roughness is a site-specific parameter. It is fitted from ERS observations acquired during the dry season with the Oh et al. (1992) model and kept constant during the wet season (Frison et al., 1998).

Because of the land surface heterogeneity, the spatial resolution of the WSC data to assimilate (~50 km) precludes from setting a priori local measurement values to the STEP model parameters. However, the WSC observations cannot be used to control every parameters of the land surface model. In particular, certain are “nonobservable” by the instrument. Here, a simple Monte Carlo uncertainty analysis is carried out to identify the sensitivity of the simulated backscattering coefficient as a function of realistic variations of the STEP model parameters. Sixteen key parameters of the STEP model are a-priori-selected for the sensitivity analysis. The implemented sensitivity method used here is a simplification of the Bastidas et al. (1999) method.

A simulated time series of the backscattering coefficient using the coupled RTM–STEP models is computed from a-priori-chosen parameter values with a temporal repetitivity that mimics the WSC instrument. This simulation is taken as our reference simulation  $\hat{\sigma}_{\text{reference}}^0(t)$ . A set of M parameter vectors  $\mathbf{p}_m$  (16-dimensional) is then generated stochastically using a uniform distribution between two extreme values either taken from ground measurements or found in the bibliography. The corresponding M simulations  $\hat{\sigma}_m^0(t, \mathbf{p}_m)$  ( $m=1..M$ ) of the RTM–STEP model runs are computed. The performance of the  $m^{\text{th}}$  simulation  $\hat{\sigma}_m^0(t, \mathbf{p}_m)$  with regards to the reference time series  $\hat{\sigma}_{\text{reference}}^0(t)$ , is evaluated by the quadratic cost function  $J_m$ :

$$J_m[\mathbf{p}_m] = \left[ \frac{1}{n} \cdot \sum_{i=1}^n (\hat{\sigma}_m^0[t_i, \mathbf{p}_m] - \hat{\sigma}_{\text{reference}}^0(t_i))^2 \right]^{\frac{1}{2}} \quad (3)$$

The whole set of simulations is then divided between acceptable solutions and nonacceptable ones. Assuming that the ERS WSC observations are contaminated by a white noise, every solution with a cost lower than twice the observation noise level is kept as acceptable (Jarlan, Mazzega et al., 2002). Cumulative parameter distributions are computed from the subset of N “acceptable” solutions and compared to cumulative distributions from the M-N “nonacceptable” ones (a priori nearly uniform apart from numerical sampling errors). If these two distributions significantly differ from each other, the parameter is assumed to have a nonnegligible influence on the backscattering coefficient.

The stochastic feature of the method precludes from a systematic exploration of the whole parameter space. However, a preliminary study shows that, apart from very local nonlinear phenomenon, 100,000 samples are sufficient to identify the general behaviour of the coupled RTM–STEP models. The number of acceptable solutions is  $N=4543$  from the 100,000 potential solutions tested within this study (thus, 4.5% of total number of function evaluation). Results show that the backscattering coefficient is sensitive to 10 parameters listed in Table 3 in decreasing order in terms of sensitivity. The sensitivity analysis is repeated with a different reference simulation and leads to similar results.

Table 3  
Selected STEP model parameters used in WSC assimilation

Symbol	Definition	Units	Range
$B_g0$	Initial above ground green biomass	$\text{g. m}^{-2}$	[0.1;3]
$C_3$	$C_3$ herb contribution	%	[0;100]
$e_{\max}$	Maximum conversion efficiency	$\text{g DM MJ}^{-1}$	[4;8]
SLA0	Specific plant area at emergence	$\text{cm}^2 \text{g}^{-1}$	[180;280]
SLAslope	Slope of the relation $\text{SLA}=\text{SLA0}*\exp(-\text{SLAslope}*t)$	–	[0.007;0.06]
$H_{\text{emergence}}$	Vegetation water content at emergence	%	[0.7;0.9]
$H_{\text{peak}}$	Vegetation water content at herbaceous mass peak	%	[0.35;0.45]
$K(0)$	Infiltration time constant for superficial layer (layer 0)	$\text{cm day}^{-1}$	[100;1200]
$a$	Parameters of the soil water	–	[3000;6000]
$b$	Resistance equation	–	[200;1200]

### 3.2. The assimilation scheme

The assimilation scheme consists of calibrating the retrievable uncertain parameter vector  $\mathbf{p}$  (10-dimensional) of the STEP model using WSC data (1) to improve the simulations of the state variables of the model (mainly, herbaceous mass and water content of the upper soil profile) and (2) to study the impact of the WSC data assimilation on simulated water fluxes. This kind of assimilation problem is called parameter identification (Delécolle et al., 1992; Evensen et al., 1998; Maas, 1988). It is usually solved by an optimization process searching for the RTM–STEP parameters that minimize the distance between ERS observations and the RTM–STEP-simulated backscattering coefficient. The distance between the  $n$  observations at time  $t_i$ ,  $\sigma^0(t_i)$  and the corresponding model prediction  $\hat{\sigma}_{\text{scene}}^0[t_i, \mathbf{p}]$  is computed using the following cost function  $J$ :

$$J[\mathbf{p}] = \left[ \frac{1}{n} \cdot \sum_{i=1}^n \frac{(\hat{\sigma}_{\text{scene}}^0[t_i, \mathbf{p}] - \sigma^0(t_i))^2}{C(t_i)} \right]^{\frac{1}{2}} \quad (4)$$

where  $C(t_i)$  is a nonstationary covariance function that “overweighs” the high backscattering observations that mainly control the simulated herbaceous mass peak. Weights are set to 1/0.1 for the backscattering peak and to 1/1 for the lower backscattering observations, with a linear variation between these two extreme values. The better the solution, the lower the cost value is. Following Jarlan et al. (2003), only those data acquired during the period of vegetation growth are kept to identify the 10 parameters.

The cost function minimization is carried out using an evolutionary algorithm (EA). EAs are stochastic optimization methods crudely mimicking Darwinian evolution: the adaptation to a given environment is the result of natural selection (survival of the fittest) and blind variations. The most well-known branches of Evolutionary Computation

(EC) are evolution strategies (ES), evolutionary programming, genetic algorithms and genetic programming (Schoenauer & Michalewicz, 1997). EAs exhibit great potentialities to solve complex real-world problems. From the point of view of optimization, EC are very powerful zeroth-order methods (that is, they only require value of the function to optimize) that can find the optimum of quite rough functions. These algorithms work on a population of individuals, each representing a search point in the space of potential solutions. The population is arbitrarily initialized and evolves towards better and better regions of the search space through the application of processes of selection, mutation, and recombination (Michalewicz, 1996; Rechenberg, 1973; Schwefel, 1981). These kind of stochastic algorithms have been already used with success within the frame of various environmental inverse problems (Fang et al., 2003; Jarlan et al., 2003; Wang, 1997).

Here, we use an evolution strategies (ES) algorithm developed by Schoenauer et al. (1995). In this particular design, the individuals of the generation at time step  $t+1$  are the best individuals among the parents plus offspring of generation at time  $t$  (the so-called ES+strategy). The number of parents and the number of offspring are set to 50. The total number of generations is set to 50. The autoadaptive mutation probability and the crossover probability are set to 0.6 and 0.4, respectively. In the data assimilation process, we shall not consider only the optimal parameter set  $\mathbf{p}^{\text{opt}}$  (i.e., that minimizes the cost function) but also the whole set of acceptable solutions. The acceptability criterion for a given trial parameter set corresponds to an associated cost function lower than twice the data noise level (Jarlan et al., 2003). Accordingly, thanks to the stochastic feature of the ES algorithm, several parameter vectors are tested, and those that fulfil the acceptability criterion are kept.

The schematic description of the assimilation procedure is given in Fig. 1. Daily outputs from the Land Surface

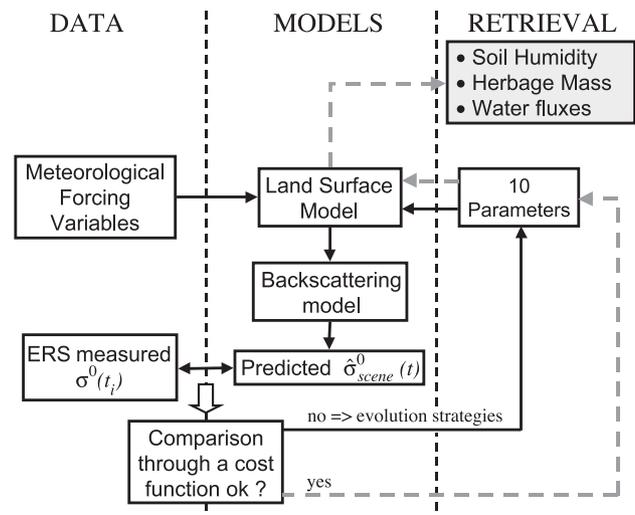


Fig. 1. Schematic description of the assimilation scheme based on the evolution strategies algorithm.

Model STEP—namely, the herbaceous mass and cover fraction, the vegetation and soil water contents—are used as inputs of the radiative transfer model to simulate a temporal series of the backscattering coefficient. These simulations are compared to WSC observations through the cost function  $J$  (Eq. (4)). If  $J$  is higher than twice the data noise level, the parameter vector is updated by the evolution strategies algorithm. If  $J$  is lower, this parameter vector is used to compute an acceptable simulation of daily herbaceous mass, soil water content and water fluxes. The algorithm is stopped when the maximum number of generations is reached.

The assimilation scheme is firstly evaluated using twin experiments. A set of backscattering observations is built using the RTM–STEP models with a known parameter vector  $p_a$ . The assimilation scheme is then applied to this artificial data set. Results show that the simulated backscattering coefficient time series is recovered with a cost lower than 0.01 dB, and the optimized parameter vector is equal to  $p_a$  (apart from the error due to the numerical sampling of the evolution strategies algorithm).

### 3.3. Running the assimilation scheme

The assimilation scheme aiming at the 10-parameter vector identification is run for the four Gourma sites during the four years: 1992, 1993, 1999 and 2000 (16 cases). Compared to the above sensitivity study performed using a Monte Carlo approach, ES use a more ‘clever’ sampling strategy of the parameter space. As a result, the number of function evaluations necessary to obtain acceptable solutions is much lower. For the 50 generation computed (about 1900 different cost evaluations), 400, on average, are kept as acceptable (21% versus 4.5% for the Monte Carlo method). The distributions of acceptable solutions are summarised by whisker plots (minimum and maximum values, first and third quartiles and median value). Simulations and ground measurements are compared by using the average absolute error (AAE).

Simulations of the RTM–STEP models after WSC data assimilation are called STEP-WSC. The STEP model is used to compute two other groups of simulations to which STEP-WSC are compared. The free runs of the model, with a set of average parameters, are called STEP-free. On the whole, these free runs without any calibration provide erroneous herbaceous mass evolution. When compared to the 21 ground measurements, the associated AAE is equal to 615 kg DM/ha. Over the reference sites where ground observations and measured soil characteristics are available, herbaceous mass measurements are used to calibrate two fitting parameters of the STEP model ( $\varepsilon_{C_{\max}}$  and  $B_g0$ ) using a look up Table approach (Mougin et al., 1995). Once the STEP model has been calibrated, it allows for a realistic simulation of herbaceous mass development with an AAE about 139 kg DM/ha. This last set of simulations, calibrated with ground measurements, are called STEP-GM.

## 4. Results and discussion

The time evolution of the STEP-WSC simulated state variables (herbaceous mass and soil moisture of the upper layer and water fluxes) is first analyzed. STEP-WSC simulations of herbaceous mass are compared to ground measurements. The associated simulated daily water fluxes are analyzed over the AMMA super site ‘‘Hombori’’ (15.3°N, 1.5°W) during the 1992, 1993, 1999 and 2000 seasons. Because of the lack of direct water fluxes observations, the temporal variability of the water fluxes after data assimilation is evaluated by a comparison with previous field campaigns measurements acquired on the same kind of ecosystems. Error bars on these simulated variables are then calculated using a stochastic approach.

### 4.1. Simulation of herbaceous mass

Fig. 2 illustrates the interannual variability of backscattering coefficient (a–d), herbaceous mass (e–h) and 15-day averaged soil volumetric humidity (i–l) over the AMMA project super site (named ‘Hombori’). With regards to climate, the long-term average of rainfall amount of the Hombori site is about 321 mm/year (over 1984–2002). With an annual report of 521 mm, year 1999 is an excellent one. Years 1993 and 2000 are close to the average, and year 1992 is quite dry (Table 4).

Time series of the simulated backscattering coefficient after assimilation and the corresponding WSC observations are drawn in Fig. 2(a–d). The fitting is only performed during the vegetation growth period as simulated by STEP-WSC. The duration and time shift of this period show a year-to-year variability (in relation with rainfall amount and distribution): from DOY 210 to DOY 275 for year 1999 and from day of the year (DOY) 190 to DOY 250 for year 1993. The agreement between observations and simulation is acceptable for the four cases. As a result of the nonstationary covariance function (see Eq. (4)), the high values of the WSC observations are better fitted than the lower ones.

The STEP-WSC herbaceous mass time series and corresponding STEP-free and STEP-GM simulations are drawn in Fig. 2(e and h) together with ground measurement data. As expected, the best fit to ground data is obtained with the STEP model calibrated with ground measurements (STEP-GM), no other constraints being added. However, the constraints brought by the WSC observations significantly improve the herbaceous mass time series when compared to the free run simulations (STEP-free). The mean parameters used to compute STEP-free simulations do not allow any vegetation development during the year 1992. For year 1999, the free run underestimates the herbaceous mass, particularly at the peak, whereas, for year 2000, the herbaceous mass is overestimated by the STEP-free simulations. By calibrating the 10 uncertain parameters, the assimilation scheme corrects for this erroneous behaviour of the model, and both the median and optimal STEP-

**AMMA project super site (Hombori, Mali, 15.3 N)**

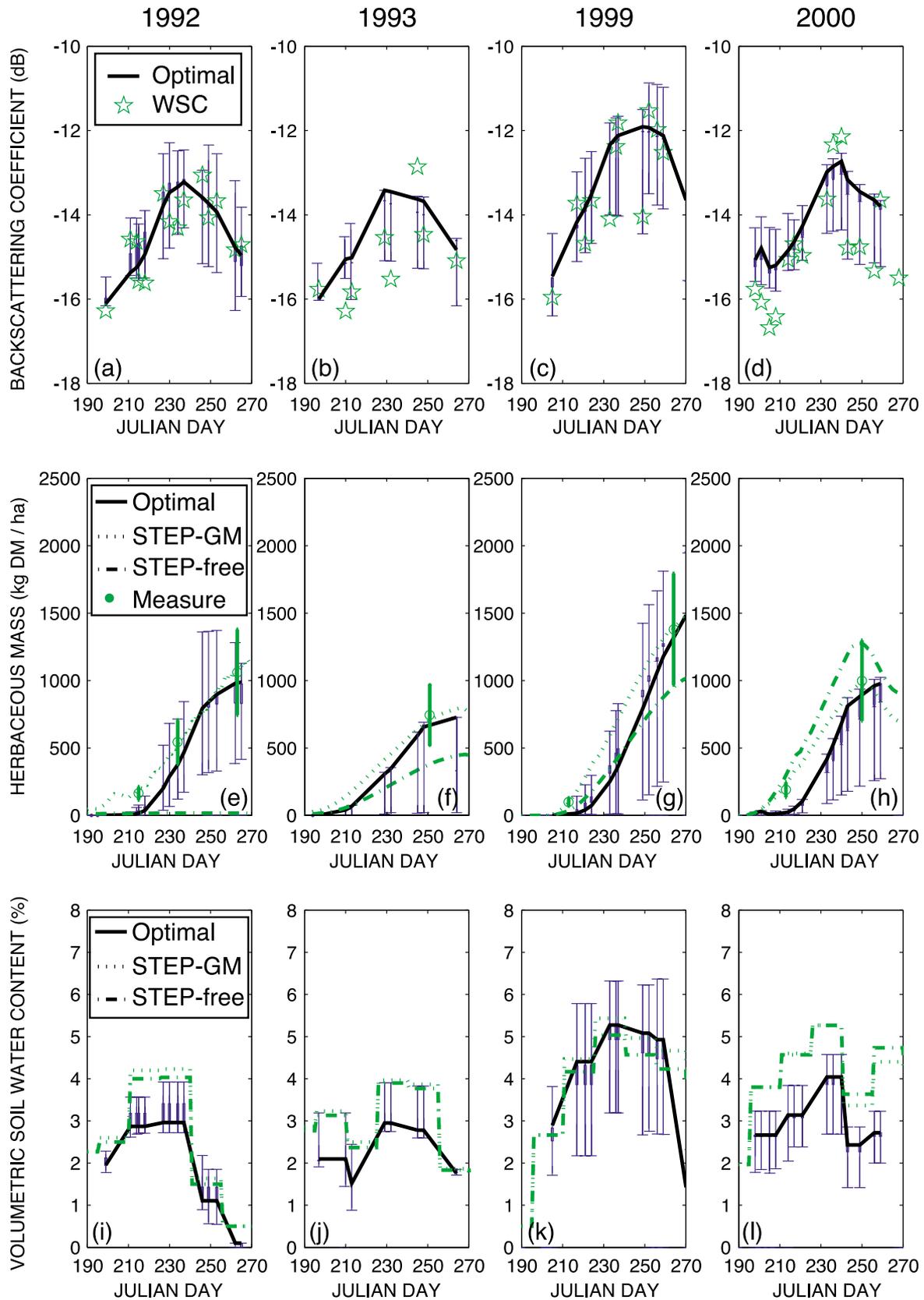


Fig. 2. Interannual variability: temporal evolution of the acceptable solutions for the Hombori site during years 1992, 1993, 1999 and 2000. Error bars represent whisker plot of the distribution. (a–d) Backscattering coefficient. (e–h) Herbaceous mass. (i–l) Soil water content of the upper soil profile. Optimal solution (—), STEP simulations of herbaceous mass without calibration (STEP-free) and calibrated with ground measurements (STEP-GM) are added for comparison.

WSC solutions are close to the ground measurements. Near the herbaceous mass peak, the absolute difference between STEP-WSC optimal solutions and ground measurements is 54, 61, 32 and 88 kg DM/ha for years 1992, 1993, 1999 and 2000, respectively. Incidentally, it should be noted that the exact date of the mass peak is not accurately known. On one hand, the temporal repetitivity of the radar sampling (1.8 data/10 days) precludes from a precise emerging and peak dates recovery. On the other hand, most in situ sites are far from any road and surveyed just when the occurrence of the mass peak is expected regionally.

At the beginning of the growing season, the simulated vegetation development starts with an average 14-day delay (16 cases considered) with STEP-WSC with regard to the STEP-GM simulations. As a consequence, ground measurements acquired at the beginning of the growing season such as that of year 2000 are strongly underestimated (Fig. 2(h)). As well, the absolute time shift of herbaceous mass peak between STEP-WSC and STEP-GM is about 9 days on average, which corresponds roughly to the WSC temporal repetitivity. Besides, from the middle stages of development to the mass peak, both herbaceous mass time series are in good agreement.

Fig. 2(i and l) show the 15-day period averaged water content of the upper soil profile. The STEP-GM and STEP-free simulations have been added for comparison. Because of the high global radiation and of the small field capacity of sandy soils, the high rate of evaporation from bare soil after a rainfall event produces a strong day-to-day variability of water content in the upper soil profile. Indeed, the surface layer completely dries out within less than 5 days (unpublished data). This rapid drying of the soil has also been observed by Wallace et al. (1993) for sandy soil under millet in a Sahelian environment.

The parameters of the water budget model (time infiltration constant and resistance to soil–atmosphere transfer) control the simulated soil moisture by modifying the “drying rate” after a rainfall event. For years 1992, 1993 and 2000, STEP-WSC 15-day averaged soil water content simulations are lower than the STEP-GM ones during the whole growing season. For year 1999, the correction applied to soil water content simulations is less important, and STEP-WSC and STEP-GM simulations are in better agreement except at the beginning of the growing season (from DOY 190 to 220). This can be mainly explained by the vegetation cover fraction difference. When the vegetation cover fraction is low (e.g., Hombori-2000), the time variations of soil water content is strongly dependent on soil evaporation which is controlled by the assimilation scheme through the soil resistance to water transfer (parameters  $a$  and  $b$  in Table 3). When the vegetation cover fraction is high (e.g., year 1999), the soil is partly masked by the vegetation, and the influence of the soil evaporation on the soil water content is less important. STEP-free and STEP-GM correspond to identical values of soil resistance to water transfer and time infiltration constant (only  $\epsilon_{\max}$

and  $B_{g0}$  differs), the slight observed differences are attributed to change in vegetation cover fraction.

In Fig. 3, the spatial variability of backscattering coefficient (a–d), herbaceous mass (e–h) and 15-day averaged soil volumetric humidity (i–l) is displayed along the north–south bioclimatic gradient of the Gourma region during year 1993 (an average year in terms of rainfall amount). From north to south, it is worth noting the regular increase of the average level of backscattering (Frison & Mougin, 1996a). The fit between simulated and observed backscattering coefficient is satisfying from the northern site (‘Rharous’) to the southern one (‘Seno’). Concerning herbage production, STEP-WSC is in better agreement with ground measurements than STEP-free one for ‘Rharous’ and ‘Hombori’ sites, STEP-free being better for ‘Gossi’ and ‘Seno.’ The simulated soil water content after WSC data assimilation for the northern (and drier) site is higher than before (STEP-free and STEP-GM). This contrasts with Fig. 2 showing a lower soil water content after WSC assimilation during dry years. This different behaviour is attributed to the longer dry season encountered in this region bordering Sahara desert and to the preceding season (1992) that was particularly dry. Soils may have been affected by a strong crusting, thus increasing the soil resistance to evaporation. When moving southward, water availability is less and less a limiting factor, allowing for a higher evaporation rate and lower soil water content than with average parameters (e.g., ‘Hombori’ and ‘Seno’ sites).

The STEP-WSC herbaceous mass simulations are then compared to all the herbaceous mass measurements (21 data distributed over four sites and four years) that are the sole available ground data at this spatial scale. Note that the STEP-WSC simulations are independent from the ground data, the optimality criterion being defined with regard to the radar data only (see Eq. (4)). On average, the optimal solutions from the set of acceptable STEP-WSC simulations are nearest to ground measurements than the median ones. In Fig. 4(a), the optimal STEP-WSC mass estimates are plotted as a function of the ground data at the date of the measurement. In Fig. 4(b), the simulated STEP-WSC mass peak estimates are plotted as a function of the corresponding simulated STEP-GM estimates. The free runs of the model (STEP-free) are not shown here as they provide poor simulations of the herbaceous mass evolution (615 kg DM/ha of AAE with regards to ground measurements date to date).

The results are satisfying on both scatter-plots with 187 kg DM/ha and 226 kg DM/ha of absolute error, respectively. This represents a noticeable improvement with regards to STEP-free. Measurements acquired at the beginning of the vegetation growth are underestimated by the assimilation results (STEP-WSC) as can be noticed in Fig. 4(a). This can be attributed to the time shift at the start of the vegetation growth between ground calibrated STEP simulations, STEP-GM and STEP-WSC. The strongest difference is associated with the southern site during the year 1993 on both plots. For

**AMMA project meso-scale site, year 1993 (Gourma region, Mali)**

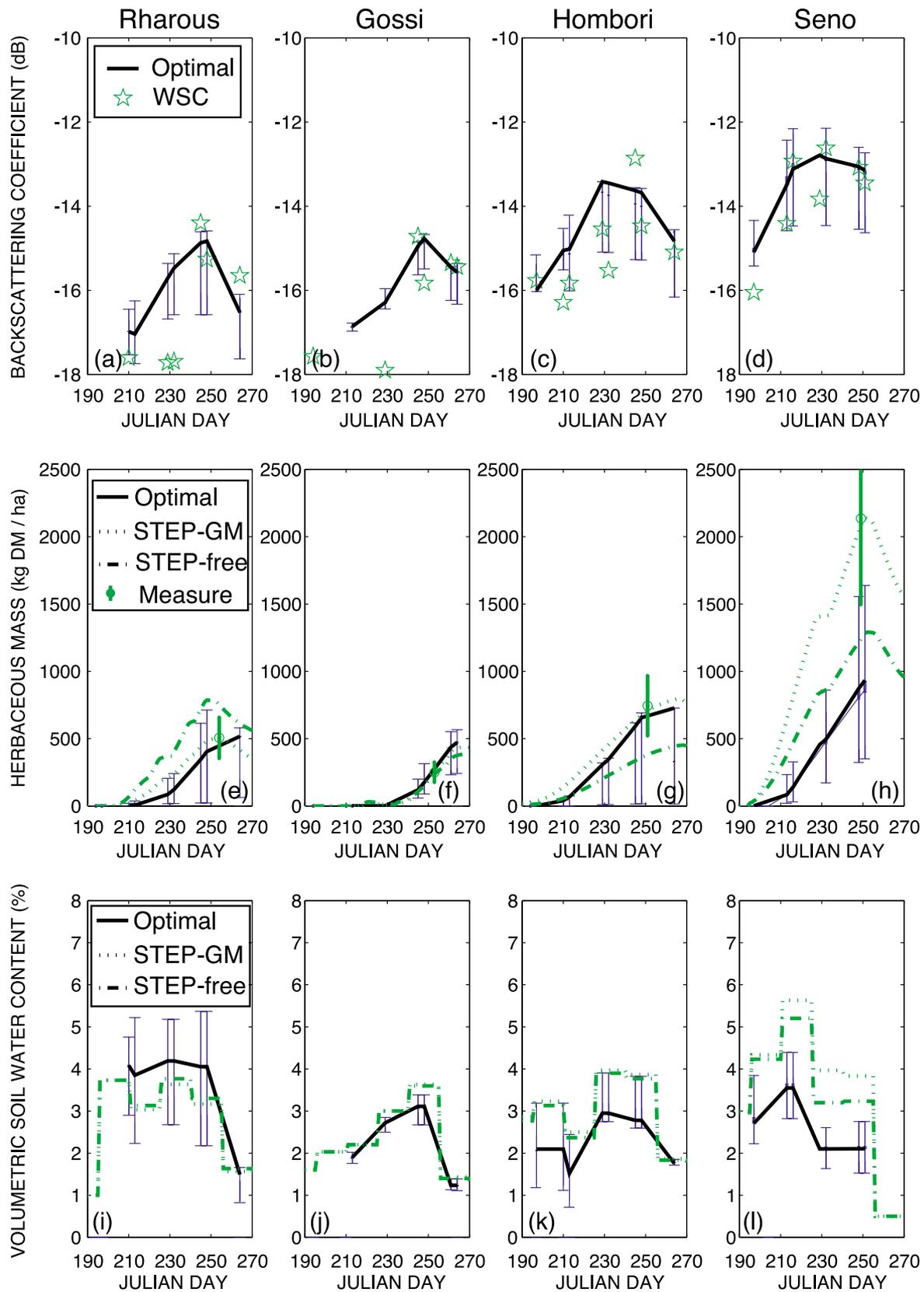


Fig. 3. Spatial variability: temporal evolution of the acceptable solutions for the year 1993 along the north–south bioclimatic gradient (‘Rharous,’ ‘Gossi,’ ‘Hombori’ and ‘Seno’ sites). Error bars represent whisker plot of the distribution. (a–d) Backscattering coefficient. (e–h) Herbaceous mass. (i–l) Soil water content of the upper soil profile. Optimal solution (–), STEP simulations of herbaceous mass without calibration (STEP-free) and calibrated with ground measurements (STEP-GM) are added for comparison.

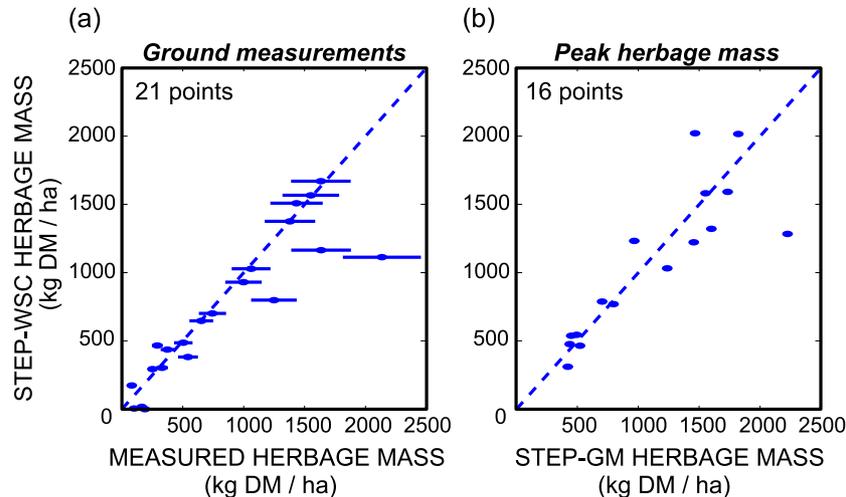


Fig. 4. Comparison of STEP simulation after WSC data assimilation with (a) ground measurements date to date and (b) simulated peak herbaceous mass by STEP calibrated on ground measurements (STEP-GM). All sites during the years 1992, 1993, 1999 and 2000 are considered. Average absolute error and correlation coefficient  $r^2$  are 187 kg DM/ha, 0.72 and 226 kg DM/ha, 0.67 for Fig. 2(a and b), respectively.

this site, the ground measurement and the STEP-GM herbaceous mass peak are equal to 2135 and 2227 kg DM/ha, respectively, whereas the STEP-WSC simulations are 929 and 1140 kg DM/ha. The maximum of herbaceous mass appears to be pretty well estimated with a 226 kg DM/ha average absolute error. These results show that the assimilation of ERS WSC data into the STEP model strongly improves its simulations of herbaceous mass with regards to the free run of the model. Furthermore, STEP-WSC compares well with STEP simulations calibrated on ground measurements, in particular, around the herbaceous mass peak.

#### 4.2. Simulations of water fluxes

We now aim at evaluating the way the assimilation of WSC data into the land surface model affects the simulated water fluxes that are critical variables for the Sahelian climate through the convection processes in the boundary layer (Nicholson, 2000). By constraining the parameters of the water and carbon submodels, the data assimilation modifies the water partition between evaporation and transpiration. This process is illustrated in Fig. 5 which shows the daily evaporation and transpiration (a) and the daily evapotranspiration (b) before and after the data assimilation (optimal solutions only) for the ‘Hombori’ site during years 1992, 1993, 1999 and 2000. Corresponding simulated herbaceous mass and soil water content are displayed in Fig. 2. The yearly rainfall and yearly statistics on daily water fluxes are given in Table 4.

Fig. 5(a) shows the time variation of soil evaporation and plant transpiration as compared to potential evapotranspiration (PET) values taken from Morel’s atlas (Morel, 1992). The primary factor governing evaporation variability over Sahel is the rainfall pattern (Sivakumar, 1990). The control of the soil resistance to water transfer by the WSC data

assimilation particularly affects the STEP model simulations at the beginning of the growing season when the distribution of rainfall events is highly irregular (e.g., years 1992, 1993 and 2000), and the near soil surface evaporation dominates evapotranspiration (Lloyd et al., 1997). On the day following a rainfall event, evaporation simulated by the STEP model after assimilation follows PET very closely during the years 1992 and 2000. On the second day, the simulated values depart significantly from PET. They are well below PET on the third day, reaching values as low as 0.3 mm/day. This simulated behaviour of soil evaporation after data assimilation is in good agreement with measurements performed during the Hapex Sahel Experiment (Kabat, Dolman et al., 1997; Wallace & Holwill, 1997). Actual evaporation may even be higher than PET as already observed by Gash et al. (1991) over a fallow savannah and by Culf et al. (1993) over a patterned woodland. On average, the data assimilation leads to higher evaporation rates than those simulated before assimilation. This higher evaporation leads to lower values (as already noticed in Fig. 2(e and f)) and stronger day-to-day variability of the soil water content, particularly, during the 1992 and 2000 rainy seasons. Afterwards, in the middle of the growing season, when vegetation cover fraction increases and rainfall distribution becomes more regular, the evaporation time variation is smoother. On year 1999, the Hombori site received particularly abundant and regular rainfalls, allowing for a high vegetation development (the simulated vegetation cover fraction at herbaceous mass peak after assimilation reaches 0.85). As a consequence, the soil was strongly masked by vegetation, and the simulated evaporation is smooth along the whole growing season.

The percentage of soil evaporation to total rainfall amount exhibits a strong interannual variability and ranges from 21% to 73% (Table 4). On average, over the four studied years, 57% of the infiltrated water is lost by soil

**AMMA project meso-scale super site (Hombori, Mali, 15.3 N)**

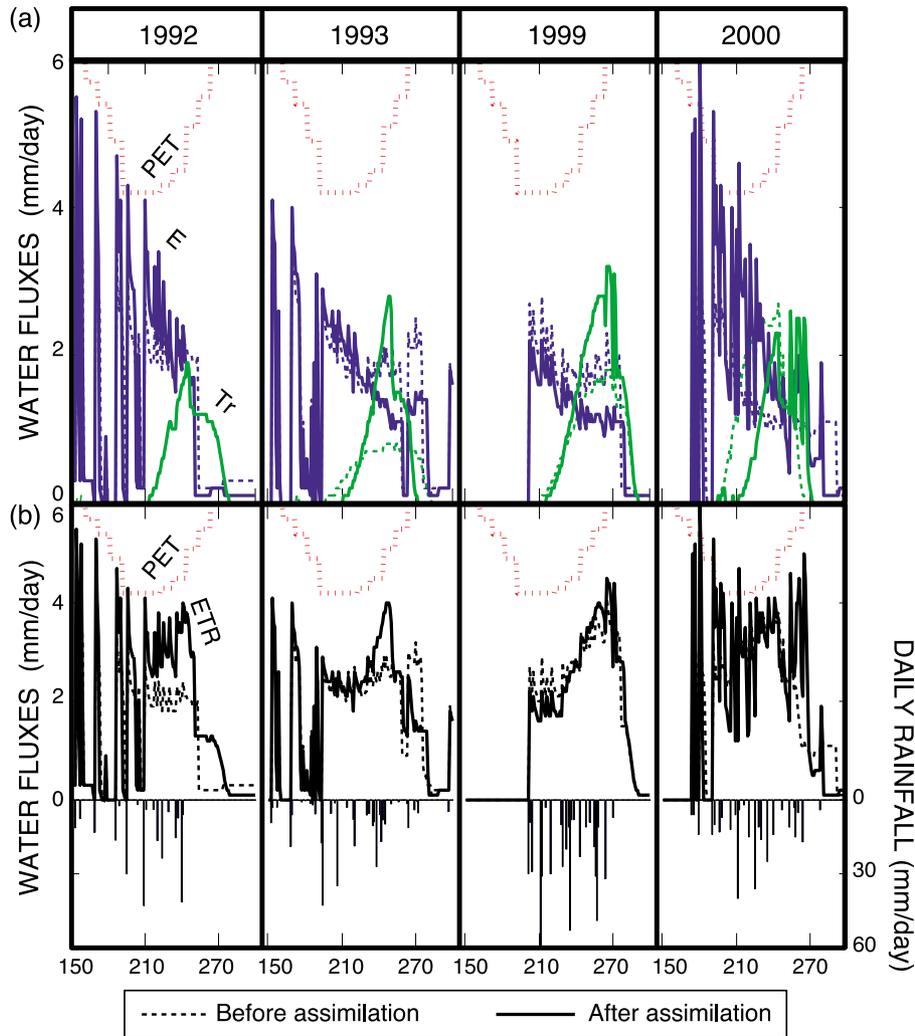


Fig. 5. Effects of WSC data assimilation on simulated daily water fluxes over the Hombori site during years 1992, 1993, 1999 and 2000. (a) Evaporation (E), transpiration (Tr) and potential evapotranspiration (PET, Morel, 1992). (b) Total evapotranspiration (ETR), potential evapotranspiration (Morel, 1992) and daily rainfall. For each figure, both simulations before (· · · · ·) and after (—) data assimilation are plotted.

evaporation. This percentage tends to increase in dry years (e.g., 73% loss during year 1992). This has also been observed by Wallace and Holwill (1997).

Differences between transpiration fluxes before and after assimilation are mainly associated to differences in simulated herbage production and thus in leaf area index (Wallace et al.,

1990, 1993). The assimilation of WSC data leads to more realistic simulations of transpiration fluxes, with higher values encountered during year of high vegetation growth. In particular, year 1999 exhibits the highest yearly transpiration rate with 115.5 mm/year (cf. Table 4) and herbage production (1400 kg DM/ha). The lowest values are obtained

Table 4

Yearly rainfall and statistics on water fluxes over the ‘Hombori’ site (15.3°N, 1.5°W) during the 1992, 1993, 1999 and 2000 growing seasons (before and after data assimilation)

	Rainfall (mm)	Evaporation (mm) (% rainfall)		Transpiration (mm) (% rainfall)		Average evapotranspiration (mm/day)	
		Before	After	Before	After	Before	After
1992	276.5	199.0 (72)	201.6 (73)	0.3 (0.1)	61.0 (22)	1.41	1.86
1993	306.1	203.4 (66)	190.8 (62)	46.3 (15)	70.3 (23)	1.82	1.90
1999	521.0	145.5 (28)	109.3 (21)	82.1 (16)	115.5 (22)	2.39	2.36
2000	306.0	140.3 (46)	219.0 (72)	105.9 (34)	75.0 (24)	2.53	3.01

for year 1992 (61 mm/year with 750 kg DM/ha). As well, the highest daily value is simulated at peak herbaceous mass for the year 1999 (3.2 mm/day at DOY 266).

Finally, WSC data assimilation allows the time variation of both evaporation and transpiration to be corrected. Higher day-to-day simulated variability of soil evaporation after assimilation is in good agreement with previous measurements available over similar ecosystems. High values close to PET at the beginning of the growing season are also in better agreement with what has been measured until now. The correction of simulated herbage production affects vegetation transpiration fluxes which exhibit, after assimilation, a more realistic behaviour with values better correlated with ground observations of herbaceous dry matter production.

#### 4.3. Uncertainty calculation

The comparison of the STEP-WSC simulations with ground measurements allows to check the overall consistency of the model physical behaviour. The parameter space semistochastic sampling done by the evolution strategies algorithm allows the error bar on every state variables of the STEP model to be estimated from the set of acceptable solutions (materialized by the whisker plots; Fig. 2(a–f)). The expected uncertainties to be associated with the herbaceous mass or water fluxes simulations depend on the WSC noise level as well as on the accuracies of the coupled RTM–STEP models and meteorological forcing variables. Error bars on the STEP simulations induced by uncertainties on the meteorological forcing variables have to be quantified. Rainfall and global radiation are the most important forcing factors as they govern both the carbon budget (through photosynthesis and water availability for the plant growth) and water budget of the model. Thus, noise is artificially added only to rainfall and to global radiation.

Rainfall in the Sahel is mostly generated by squall lines, which typically arrive at 3-day intervals throughout the rainy season. However, storms within these squall lines are convective, and the spatial distribution of rainfall is thus highly variable, in particular, at the scale of the WSC resolution cell. Obviously, this spatial variability cannot be taken into account with the rainfall data obtained from one ground station available per site. Consequently, the rainfall data used in the assimilation procedure are probably not representative of the resolution cell. As the water content of the first soil centimetres is only slightly influenced by the rainfall amount because of a low field capacity, only the rainfall time distribution is perturbed. To this end, the number of rainfall events and the amount of each event per decade are stochastically modified in such a way that this new distribution follows the average statistics obtained from our rainfall database (Jarlan et al., 2003). Moreover, the global radiation from the Morel (1992) atlas is contaminated by a 30% white noise (uniform distribution around atlas value), which is the typical noise level that can be expected

for estimates derived from a remote sensing instrument such as METEOSAT (Ba et al., 2001). The assimilation scheme is run 500 times using different noisy forcing variables, and optimal herbaceous mass and water flux time series are kept.

The uncertainty calculations are only illustrated for the Hombori-2000 and Seno-1999 cases because of computing time constraints and because of their different behaviour. Fig. 6(a–f) show the distributions of the herbaceous mass peak values for Seno-1999 and Hombori-2000 resulting from noise contamination of rainfall, global radiation and rainfall plus global radiation, respectively. The resulting distributions are likely to be quasinormal (only a much larger set of simulations would allow to test this assumption). Because of a stronger sensitivity to rainfall than to global radiation, the herbaceous mass distributions are narrower when noise is only added to the global radiation than to the rainfall (Seno-1999: 139 kg DM/ha versus 268 kg DM/ha standard deviations (S.D.), respectively; Hombori-2000: 124 kg DM/ha versus 236 kg DM/ha S.D., respectively). With regards to ground measurements, adding noise to rainfall recenters the herbage distribution on the Seno-1999 ground data and leads to a less important underestimation for Hombori-2000. Contaminating both forcing variables does not significantly change the mass distribution S.D. over the sole rainfall contamination case. This confirms the dominant sensitivity of the STEP model simulations to the rainfall distribution which is the main climatic factor governing the surface functioning over the Sahel. With both rainfall and global radiation noise contaminations, the variation coefficients (S.D. divided by the mean) of the resulting herbaceous mass distributions (Fig. 5(c and f)) are about 15.1% for Seno-1999 and 15.9% for Hombori-2000.

These uncertainty estimates on the herbaceous mass after assimilation compare well with those from Jarlan et al. (2003) based on the inversion of the radiative transfer model. For this later case, the herbaceous mass S.D. resulting from forcing uncertainties was between 230 and 271 kg DM/ha over the whole Gourma region. Relying on the land surface model, the assimilation approach appears to bring better constraints to the herbaceous mass estimation. Indeed, the standard deviations of the distribution exhibit more variability (Hombori-2000: 139 DM/ha; Seno-1999: 268 kg DM/ha) but still have values lower than those estimated with the inversion process.

Yearly, water flux distributions (evaporation, transpiration and evapotranspiration) are shown in Fig. 7(a–f) for the Seno-1999 and Hombori-2000 cases. These distributions result from the same noise contamination process of rainfall and global radiation data (see above). Mean values of yearly evaporation (Seno-1999: 223 mm/y; Hombori-2000: 220 mm/y) are higher than yearly transpiration ones (Seno-1999: 71 mm/y; Hombori-2000: 108 mm/y). As a result, the relative accuracies on yearly evaporation are lower than those of yearly transpiration: 8.1% and 28.1% for Seno-1999; 11.8% and 24.1% for Hombori-2000. The standard deviation of evapotranspiration (about 15 mm/y for both Seno-1999 and

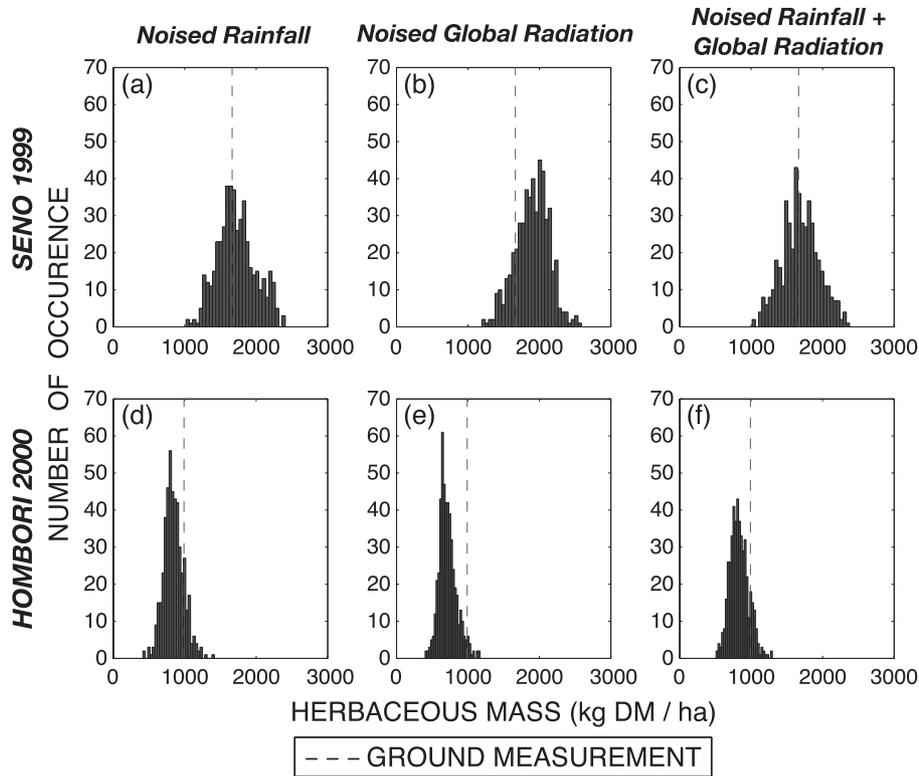


Fig. 6. Meteorological forcing variable uncertainties propagated on simulated herbaceous mass after assimilation. Panels (a–c) and (d–f) show the Seno-1999 and the Hombori-2000 cases after noising rainfall, global radiation and rainfall+global radiation, respectively. Herbaceous mass measurements are also plotted (–). Mean, standard deviation (in kg DM/ha) and variation coefficient (standard deviation/mean) are equal to 1723, 268, 15.5%; 1901, 236, 12.4%; 1691, 257, 15.2%; 844, 139, 16.5%; 716, 124, 17.3% and 831, 132, 15.9% for panels (a–f), respectively.

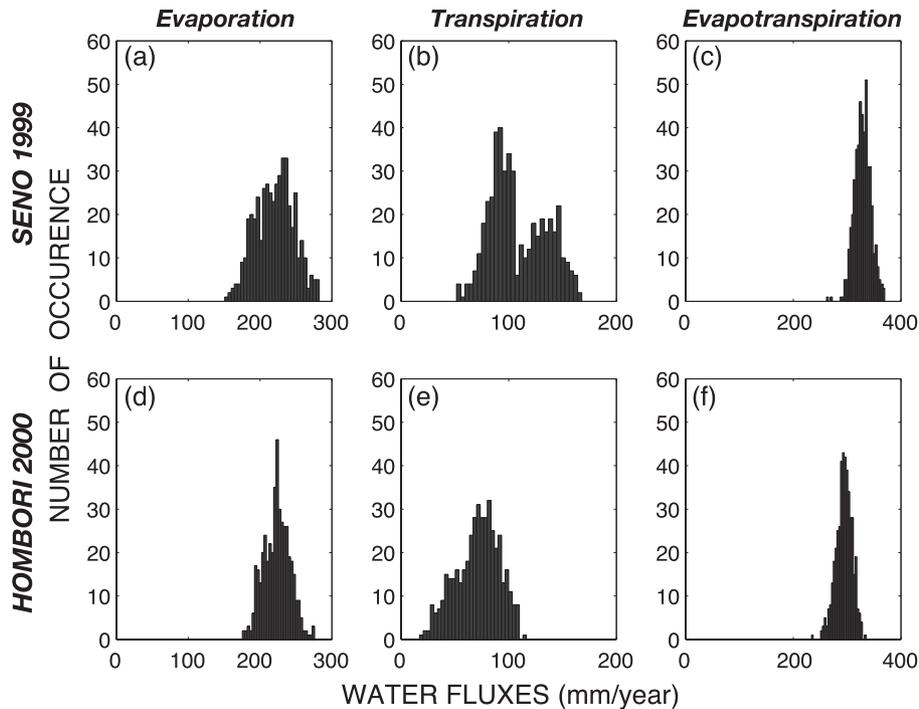


Fig. 7. Rainfall and global radiation uncertainties propagated on simulated yearly water fluxes. Panels (a–c) and (d–f) shows Seno-1999 and Hombori-2000 distributions of evaporation, transpiration and evapotranspiration, respectively. Mean, standard deviation (in mm/year) and variation coefficient (standard deviation/mean) are equal to 220, 26, 11.8%; 108, 26, 24.1%; 328, 15, 4.6%; 223, 18, 8.1%; 71, 20, 28.1% and 294, 15, 5.1% for panels (a–f), respectively.

Hombori-2000) is lower than both those of evaporation and transpiration. An interesting feature is that the monomodal herbaceous mass distributions (Fig. 6(c)) leads to a bimodal yearly transpiration distribution (Fig. 7(b)) for Seno-1999. This bimodality highlights the non linear behaviour of the STEP model. After analysis, the left mode of yearly transpiration distribution (low transpiration rates) appears to be associated with low herbaceous mass development and high evaporation rate. In fact, the assimilation algorithm chooses higher yearly evaporation rate to compensate for low yearly transpiration so that the total evaporation remains more or less constant. This in turn modifies the partition between the simulated soil evaporation and plant transpiration (via the land surface model) and can change the atmospheric dynamics when coupled to a mesoscale atmospheric dynamical model. This underdetermination illustrates the danger of using only one data set during the assimilation process for controlling a model characterized by nonlinear equations such as the STEP model.

## 5. Conclusion and perspectives

The objective of this study is to correct the simulated herbaceous mass, soil moisture and water fluxes derived from a simple land surface model through the assimilation of coarse remote sensing radar data over Sahel. Thanks to a stochastic sensitivity study, the parameters of the land surface model that can be determined with remote sensing observations are identified. The assimilation scheme relies on an evolution strategies algorithm (a semistochastic parameter identification method). This kind of approach is particularly well suited for this complex (nonlinear) inverse problem. The resulting time variation of the herbaceous mass is analyzed and compared to ground measurements performed on four Sahelian sites over four years. The time variations of the simulated standing herbaceous mass after remote sensing data assimilation exhibit a 187 kg DM/ha error which compares well with a previous study of Jarlan et al. (2003) based on the inversion of the radiative transfer model.

In contrast, with a repetitivity of 1.8 data/decade, the assimilation of wind scatterometer data does not significantly affect the short time scale dynamics of the upper soil moisture. The effects of data assimilation on water fluxes are analyzed and compared with available measurements made by various authors on the same kind of ecosystem. Because of a lack of direct observations of water fluxes on the considered study sites, only the time behaviour and absolute value consistency are checked. After the data assimilation, the simulated water fluxes time variations exhibit a higher day-to-day variability and higher values close to the potential evapotranspiration. This behaviour after WSC assimilation is in better agreement with those previous ground campaigns.

Finally, uncertainties on the herbaceous mass and water fluxes resulting from uncertainties on climatic forcing variables (rainfall and global radiation) are estimated using

a stochastic approach. The uncertainty ranges between 139 and 268 kg DM/ha for the herbaceous mass. This work shows that the accuracy of the rainfall data strongly conditions the uncertainty level associated with the simulated herbaceous mass after WSC data assimilation. Nevertheless, the assimilation of WSC data appears to be well suited to control the carbon budget of our land surface model. Concerning the water fluxes uncertainties, the estimates are less robust (in particular, on yearly transpiration) and must be improved in a next study by using additional information from satellite remote sensing at different wavelengths. Furthermore, water fluxes measurements by eddy correlation techniques are already planned within the frame of the AMMA project during the 2004 growing season and will allow for a further validation of the approach.

## Appendix A. Main equations of the STEP model

The STEP model is composed of two submodels: carbon budget submodel (vegetation growth) that simulates the main processes of the vegetation development (photosynthesis, allocation, respiration and senescence) and the water budget submodel that simulates the soil water dynamic (runoff, percolation and drainage) and the exchanges of water with the atmosphere (water lost by soil evaporation and by plant transpiration). Modifications from the previous version (Mougin et al., 1995) are detailed below.

### A.1. Carbon budget

The dynamics of the green above-ground mass  $B_g$  and green mass root  $B_r$  are described using a set of two differential equations:

$$\frac{dB_g}{dt} = \alpha_1 \cdot a \cdot PS_n + \alpha_2 \cdot B_g \quad (A1)$$

$$\frac{dB_r}{dt} = \alpha_3 \cdot (1 - a) \cdot PS_n + \alpha_4 \cdot B_r \quad (A2)$$

$$PS_n = P_g - R_t \quad (A3)$$

where  $PS_n$  is the net photosynthesis calculated as the difference between gross photosynthesis  $P_g$  and respiration losses  $R_t$ ,  $\alpha_i$  are empirical parameters,  $\alpha_1 \cdot B_g$  and  $\alpha_4 \cdot B_r$  represent losses of above ground and root mass, respectively, due to vegetation senescence, and  $a$ , the fraction of allocated above-ground part, is calculated such that:

$$\frac{B_r}{B_g} = \frac{1.2}{2 + 0.01 \cdot B_g} \quad (A4)$$

The growth photosynthesis is expressed as

$$P_g = PAR \cdot \varepsilon_i \cdot \varepsilon_{Cmax} \cdot f(\Psi1) \cdot f(TI) \quad (A5)$$

Photosynthesis depends on the intercepted photosynthetic active radiation  $PAR_i$ , a maximum energy-conversion efficiency,  $\varepsilon_{\max}$ , one function of leaf water potential expressing the effect of water limitation  $f(\Psi)$ , and one function expressing the effect of temperature limitation  $f(T)$ .  $PAR \cdot \varepsilon_i = PAR_i$  is the photosynthetically active radiation (part of the incoming global radiation between 400 and 700  $\mu\text{m}$ ) that is intercepted by plants.  $\varepsilon_i$  is linked to herbaceous mass leaf area index (area of leaf by  $\text{m}^2$  of soil).  $\varepsilon_i$  and  $\varepsilon_{\text{Cmax}}$  vary between 0 and 1.

The green  $LAI_g$  is computed from the green herbaceous mass  $B_g$  through the specific leaf area  $SLA_g$  which depends on the value at plant emergence  $SLA_0$  and on the slope of the relation  $SLA_g$  versus time  $SLA_g$ slope:

$$LAI_g = SLA_g \cdot B_g$$

$$SLA_g = SLA_0 \cdot e^{-SLA_g \text{Slope} \cdot t} \quad (\text{A6})$$

Typical values of  $SLA_0$  and  $SLA_g$ slope are calculated from ground measurements (Tracol, unpublished data).

Growth starts at plant emergence the date of which is calculated from the soil moisture content of the first five centimetres. If the soil moisture content of the upper layer stay above the wilting point (calculated from empirical relationship with the soil characteristics), seeds emerge, and vegetation growth begins. The initial value of the above ground mass is set to  $B_{g0}$ . Typical values comprised between 0.1 and 3 kg MS/ha.

The present vegetation is a mixture of two kinds of herbs determined by their photosynthesis type ( $C_4$  and  $C_3$ ). The percentage of the  $C_3$  grasses must be known. This percentage is highly variable from one year to another.

The total vegetation cover fraction  $V_{\text{cft}}$  is computed from the leaf area index  $LAI$  and the canopy extinction coefficient  $k_c$  as

$$V_{\text{cft}} = 1 - e^{-k_c \cdot LAI} \quad (\text{A7})$$

The model also calculates the time evolution of vegetation gravimetric moisture content needed for the radar radiative transfer model. Two fitting parameters are needed to calculate this evolution: the gravimetric water content at emergence  $H_{\text{emergence}}$  and the gravimetric water content at herbaceous mass peak  $H_{\text{peak}}$ . Typical values are given in Guerin et al. (1991).  $H_p$  is maximum at emergence at around 80% and is around 40% at herbage peak.

## A.2. Water budget

The distribution on the soil profile is simulated according to a ‘‘tipping-bucket’’ scheme: water that penetrates the soil is assumed to fill the consecutive soil layers up to field capacity  $FC_i$ . The soil is composed of

four layers. The soil water content of layer  $i$ ,  $H_v(i)$ , is simulated following this set of differential equations:

$$\frac{dH_v(1)}{dt} = P - R - E_1 - D_1 \quad (\text{A8})$$

$$\frac{dH_v(i)}{dt} = D_{i-1} - E_i - Tr_i - D_i \quad (\text{A9})$$

$$\text{and if } W_i > FC_i D_i = (D_{i-1} - FC_i) / K_i$$

$$\text{else } D_i = 0 \quad (\text{A10})$$

where  $P$ ,  $R$ ,  $E_i$ ,  $D_i$  and  $Tr_i$  denote precipitation, run-off, evaporation of layer  $i$ , drainage of layer  $i$  and contribution to transpiration of layer  $i$ , respectively.  $K_i$  represents the infiltration time constant of layer  $i$ .

Plant transpiration  $Tr$  and soil evaporation  $E$  are computed following the Penman–Monteith equation (Monteith, 1965). For this, the soil and canopy resistance to water vapour transfer must be known. The soil resistance  $r_{\text{ss}}$  is calculated following Camillo and Gurney (1986) from the water content of the upper soil layer  $W_1$ :

$$r_{\text{ss}} = a \cdot (W_{\text{sat}} - W_1) - b \quad (\text{A11})$$

with  $W_{\text{sat}}$  as the soil water content at saturation and  $a$  and  $b$ , two empirical coefficients. The canopy resistance is estimated as in Mougin et al. (1995).

## References

- African Monsoon Multidisciplinary Analysis (AMMA). (2001). White book: La mousson africaine et ses composantes, 75 pp. (<http://medias.obs-mip.fr/amma>).
- Ba, B. M., Frouin, R., Nicholson, E. S., & Dedieu, G. (2001). Satellite-derived surface radiation budget over the African Continent: Part I. Estimation of downward irradiance and albedo. *Journal of Climate*, 14, 45–58.
- Bastidas, L. A., Gupta, H. V., Sorooshian, S., Shuttleworth, W. J., & Yang, Z. L. (1999). Sensitivity analysis of land surface scheme using multicriteria methods. *Journal of Geophysical Research*, 104(16), 19481–19490.
- Boulet, G., Kalma, J. D., Kerr, Y. H., & Chehbouni, A. (2002). Deriving catchment-scale water and energy balance parameters using data assimilation based on Extended Kalman Filtering. *Hydrological Sciences Journal*, 47(3), 449–467.
- Brubaker, K. L. D., Entekhabi, D., & Abd Eagleson, P. S. (1993). Estimation of continental precipitation recycling. *Journal of Climate*, 6(6), 1077–1089.
- Camillo, P. J., & Gurney, R. J. (1986). A resistance parameter for bare soil evaporation models. *Soil Sciences*, 141, 95–105.
- Cayrol, P., Kergoat, L., Moulin, S., Dedieu, G., & Chehbouni, A. (2000). Calibrating a coupled SVAT/vegetation growth model with remotely sensed reflectance and surface temperature. A case study for the HAPEX-Sahel grasslands sites. *Journal of Applied Meteorology*, 39, 2452–2472.
- Crow, W. T., & Wood, E. F. (2003). The assimilation of remotely sensed soil brightness temperatures imagery into a land surface model using Ensemble Kalman Filtering: A case study based on ESTAR measurements during SGP97. *Advances in Water Resources*, 26, 137–149.

- Culf, A. D., Allen, S. J., Gash, J. H. C., Lloyd, C. R., & Wallace, J. S. (1993). Energy and water budgets of an area of patterned woodland in the Sahel. *Agricultural and Forest Meteorology*, 66, 65–80.
- Delécolle, R., Maas, S. J., Guérif, M., & Baret, F. (1992). Remote sensing and crop production models. Present trends, ISPRS Phtogramm. *Remote Sensing*, 47, 145–161.
- Dickinson, R. E. (1995). Land processes in climate models. *Remote Sensing of Environment*, 51, 27–38.
- Evensen, G., Dee, D., & Schroter, J. (1998). Parameter estimation in dynamical models, NATO ASI. In E. P. Chassignet, & J. Verron (Eds.), *Ocean Modelling and Parameterizations*.
- Fang, H., Liang, S., & Kuusk, A. (2003). Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing of Environment*, 85, 257–270.
- Frison, P. L., & Mougou, E. (1996a). Monitoring global vegetation dynamics with ERS-1 wind scatterometer data. *International Journal of Remote Sensing*, 17(16), 3201–3218.
- Frison, P. L., & Mougou, E. (1996b). Use of ERS-1 wind scatterometer data over land surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 1–11.
- Frison, P. L., Mougou, E., & Hiernaux, P. (1998). Observations and interpretation of seasonal ERS-1 wind scatterometer data over northern Sahel (Mali). *Remote Sensing of Environment*, 69(3), 233–242.
- Gash, J. H. C., Wallace, J. S., Lloyd, C. R., Dolman, A. J., Sivakumar, M. V. K., & Renard, C. (1991). Measurements of evaporation from fallow Sahelian savannah at the start of the dry season. *Quarterly Journal of the Royal Meteorological Society*, 117, 749–760.
- Guerin, H., Friot, D., Mbaye N., & Richard, D. (1991). Alimentation des ruminants Domestiques sur pâturages Naturels Sahéliens et Sahelo-soudanien: Etude Méthologique dans la région du Ferlo au Sénégal, Institut d'Élevage et de Médecine vétérinaire des Pays Tropicaux (I.E.M.V.T.), 115 pp.
- Hallikainen, M. T., Ulaby, F. T., Dobson, M. C., El-Rayes, M. A., & Wu, L. -K. (1985). Microwave dielectric behavior of wet soil: Part I. Empirical models and experimental observations. *IEEE Transactions on Geoscience and Remote Sensing*, 23, 25–34.
- Hiernaux, P. H. Y., & Justice, C. O. (1986). Suivi du développement végétal au cours de l'été 1984 dans le Sahel Malien. *International Journal of Remote Sensing*, 11, 1515–1531.
- Jarlan, L., Mazzega, P., & Mougou, E. (2002). Retrieval of land surface parameters from ERS scatterometer data: A brute-force method. *IEEE Transactions on Geoscience and Remote Sensing*, 40(9), 2056–2062.
- Jarlan, L., Mougou, E., Frison, P. L., Mazzega, P., & Hiernaux, P. (2002). Analysis of ERS wind scatterometer time series over Sahel (Mali). *Remote Sensing of Environment*, 81, 404–415.
- Jarlan, L., Mazzega, P., Mougou, E., Schoenauer, M., Lavenue, F., Marty, G., et al. (2003). Mapping of Sahelian vegetation parameters from ERS scatterometer data with an evolution strategie algorithm. *Remote Sensing of Environment*, 87, 72–84.
- Kabat, P., Dolman, A. J., & Elbers, J. A. (1997). Evaporation, sensible heat and canopy conductance of fallow savannah and patterned woodland in the Sahel. *Journal of Hydrology*, 188/189, 494–515.
- Kabat, P., Hutjes, R. W. A., & Feddes, R. A. (1997). The scaling characteristics of soil parameters: from plot scale heterogeneity to subgrid parameterization. *Journal of Hydrology*, 190, 363–396.
- Le Houérou, H. N. (1989). The grazing land ecosystems of the African Sahel. *Ecological Studies*, vol. 75. Berlin: Springer-Verlag.
- Lloyd, C. R., Bessemoulin, P., Cropley, F., Culf, A. D., Dolman, A. J., Elbers, J. A., et al. (1997). An intercomparison of surface flux measurements during HAPEX-Sahel. *Journal of Hydrology*, 188–189.
- Lo Seen, D., Mougou, 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 Sensing of Environment*, 52, 194–206.
- Maas, S. J. (1988). Using satellite data to improve model estimates of crop yield. *Agronomy Journal*, 80, 655–662.
- Mahrt, L. (2000). Surface heterogeneity and vertical structure of the boundary layer. *Boundary-Layer Meteorology*, 96, 33–62.
- Magagi, R. D., & Kerr, Y. H. (1997). Characterization of surface parameters over arid and semi-arid areas by use of ERS-1 wind-scatterometer. *Remote Sensing Reviews*, 15, 133–155.
- Manabé, S. (1969). Climate and the Ocean circulation: 1. The atmospheric circulation and the hydrology of the earth's surface. *Monthly Weather Review*, 97, 739–774.
- Michalewicz, Z. (1996). Genetic Algorithms+Data Structure=Evolution Programs. (Third revised and Extended Edition) Springer.
- Monteith, J. L. (1965). Evaporation and environment. *State and Movement of Water in Living Organisms (19th Symposium of the Society for Experimental Biology)* (pp. 205–234). United Kingdom: Cambridge University Press.
- Morel, R. (1992). Atlas agroclimatique des pays de la zone du CILSS. Programme Agrhymet, coopération française, CILSS, PNUD, OMM Niamey, 11 vol.
- Mougou, E., Lo Seen, D., Rambal, S., Gaston, A., & Hiernaux, P. (1995). A regional Sahelian grassland model to be coupled with multispectral satellite data: I. Description and validation. *Remote Sensing of Environment*, 52, 181–193.
- Nicholson, S. E. (2000). Land surface processes and Sahel climate. *Reviews of Geophysics*, 38, 117–139.
- Oh, Y., Sarabandi, K., & Ulaby, F. T. (1992). An empirical model and inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 30, 370–381.
- Polcher, J., Laval, K., Dümenil, L., Lean, J., & Rowntree, P. R. (1995). Comparing three land surface schemes used in general circulation models. *Journal of Hydrology*, 180, 373–394.
- Rechenberg, I. (1973). Evolution strategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution. Stuttgart: Fromman-Holzboog Verlag.
- Reichle, R., Mc Laughlin, D. B., & Entekhabi, D. (2001). Variational data assimilation of microwave radiobrightness observations for land surface hydrologic applications. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8), 1708–1718.
- Reichle, R., Mc Laughlin, B. D., & Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble Kalman Filter. *Monthly Weather Review*, 130(1), 103–114.
- Schoenauer et al. (1995). EvolC User Manual, Technical Report, EEAAX, Ecole Polytechnique, <http://www.eark.polytechnique.fr/EvolC.html>
- Schoenauer, M., & Michalewicz, Z. (1997). Evolution computation. *Control and Cybernetics*, 26(3), 307–338.
- Schwefel, H. P. (1981). Numerical optimization of computer models. Wiley: Chichester.
- Sellers, P. J., Heiser, M. D., Hall, F. G., Verma, S. B., Desjardins, R. L., Schuepp, P. M., et al. (1997). The impact of using area-averaged land surface properties—topography, vegetation condition, soil wetness—in calculations of intermediate scale (approximately 10 km<sup>2</sup>) surface-atmosphere heat and moisture fluxes. *Journal of Hydrology*, 190, 269–301.
- Shukla, J., & Mintz, Y. (1982). Influence of land-surface evapotranspiration on the earth's climate. *Science*, 215, 1498–1500.
- 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, 321–332.
- Taylor, C. M., & Lebel, T. (1998). Observational evidence of persistent convective-scale rainfall patterns. *Monthly Weather Review*, 126, 1597–1607.
- Tracol, Y., Mougou, E., & Hiernaux, P. (2004). Testing a Sahelian grassland model against herbaceous mass measurements. *Ecological Modelling* (submitted to).
- Ulaby, F. T., & El-Rayes, M. A. (1987). Microwave dielectric spectrum of vegetation: Part II. Dual-dispersion model. *IEEE Transactions on Geoscience and Remote Sensing*, 25, 550–557.
- Wagner, W., & Scipal, K. (2000). Large scale soil moisture mapping in Western Africa using the ERS scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 38, 1777–1782.

- Wallace, J. S., & Holwill, C. J. (1997). Soil evaporation from Tiger-bush in south-west Niger. *Journal of Hydrology*, 188–189, 426–442.
- Wallace, J. S., Lloyd, C. R., & Sivakumar, M. V. K. (1993). Measurements of soil, plant and total evaporation from millet over Niger. *Agricultural and Forest Meteorology*, 63, 149–169.
- Wallace, J. S., Roberts, J. M., & Sivakumar, M. V. K. (1990). The estimation of transpiration from sparse dryland millet using stomatal conductance and vegetation area indices. *Agricultural and Forest Meteorology*, 51, 35–49.
- Wang, Q. J. (1997). Using genetic algorithms to optimise model parameters. *Environmental Modelling and Software*, No. 1, vol. 12 (pp. 27–34).
- Woodhouse, I. H., & Hoekman, D. H. (2000). Determining land surface parameters from the ERS-1 wind-scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 38, 126–140.