Tellus (1999), 51A, 453–468
Printed in UK – all rights reserved



TELLUS
ISSN 0280–6495

# Variational surface analysis from screen level atmospheric parameters

By F. BOUYSSEL*, V. CASSÉ and J. PAILLEUX, *Météo-France/CNRM, 42 Avenue G. Coriolis, 31057 Toulouse, France*

(Manuscript received 25 June 1998; in final form 9 November 1998)

ABSTRACT

A variational method to retrieve soil parameters from screen level parameters is investigated. A one-dimensional model including vertical diffusion and land-surface schemes is used. A sensitivity study evaluates the influence of soil moisture and temperature in 2 layers on the surface fluxes and the 2*m* parameters. The mean soil water content appears to be the most important soil parameter for modelling surface fluxes and 2*m* parameters. During stable conditions, the mean soil temperature is also important for a good forecast of screen level parameters. The screen-level observations are found to be informative on soil moistures during the day and on soil temperatures during the night. The characteristics of the variational assimilation method are presented in the framework of identical twin experiments. The general shape of the cost function with a 24-h assimilation period reveals the ambiguity in the repartition of water in the two soil reservoirs and the presence of secondary minima. The ability of this method to initialize soil parameters is tested with various soil water contents and vegetation fraction covers. Finally, some results show the influence of the frequency of the observations and the length of the assimilation period on the quality of the analysis.

## 1. Introduction

The surface fluxes (sensible and latent heat) may have an important influence on the boundary layer development and therefore on the general atmospheric circulation. To improve the modelling of these fluxes in numerical weather prediction (NWP) models advanced land-surface schemes have been developed during the last 10 years in meteorological centres. These schemes represent the most important soil processes and have been validated on several field experiments. However, because of their high sensitivity to soil parameters some problems arise when they are implemented in NWP models. Furthermore, some soil parameters have a long time scale evolution (exceeding weeks for the mean soil water content). An erroneous initialization of these variables will cause a poor evaluation of surface fluxes during several weeks and perhaps more because of some positive feedbacks between the surface and the atmosphere. It is therefore important to associate an assimilation scheme of surface fields when implementing a land-surface scheme.

The major problem of land surface assimilation is the lack of routine observations of soil moisture and soil temperature. These surface variables should be analysed indirectly by assimilation of other observation types. The assimilation methods currently developed can be ranged in three main classes depending on the observations used: screen level observations, rainfall rates (Jones et al., 1995; Macpherson, 1996) and remote sensing observations (infrared, microwave) (see reviews in Paloscia, 1996).

* Corresponding author.

Mahfouf (1991) showed that in some meteorological conditions 2*m* observations of temperature and relative humidity can be used to estimate soil water content. He proposed two methods: a sequential and a variational technique. The sequential approach is an optimal interpolation analysis transferring the 2*m* analysis increments into the soil variables. A surface analysis using this first technique is now in operation at Météo-France (Giard and Bazile, 1997). The variational method determines the initial state which best fits a background and the observations during a period of time. This method is potentially superior to the first one because it takes into account the non-linearities between the observations and the surface variables through the model prognostic equations. Moreover, the temporal distribution of observations is taken into account. In both cases a good 2*m* analysis is needed at each grid point (Navascues, 1997).

## 2. Description of the 1D-model

### 2.1. General characteristics

In the surface analysis presented here, the correlations between soil quantities in different model grid boxes are assumed to be small. Therefore, a surface analysis at each grid point is carried out separately. A global 4D-Var approach with the analysis of both atmospheric and surface prognostic variables would be much more difficult. Furthermore the assimilation periods are not necessarily the same, the maximum is currently six hours in the first case because of computational cost and it can be one day or more for surface to catch the diurnal cycle.

A reduced one-dimensional version of the 3D operational model ARPEGE of Météo-France is used. The physical parametrizations used are the original vertical diffusion scheme (Geleyn et al., 1995) and the land-surface scheme ISBA (Noilhan and Planton, 1989; Mahfouf and Noilhan, 1996). The vertical and surface exchange coefficients for heat and momentum are computed according to Louis (1979) and Louis et al. (1981). The interpolation at screen level height is described in Geleyn (1988). The solar and thermal radiation fluxes are considered to be independent forcing terms, and are taken from a 3D host model. All the other parametrizations (convection, precipitations, horizontal advection, ...) are represented by forcing terms on atmospheric prognostic variables. For a given grid point they are calculated as the difference between the 3D and the reduced 1D model.

### 2.2. The land-surface scheme ISBA

The land-surface scheme ISBA describes the interactions between soil, vegetation and atmosphere. It includes the treatment of soil heat content, soil water content, water interception by vegetation and aerodynamic transfer processes in the atmospheric surface layer. The scheme uses the force restore model for soil heat and water content (Blackadar, 1976; Deardorff, 1977) and the so-called $\alpha$-method for evaporation (Mahfouf and Noilhan, 1991). The scheme includes a representation of snow (Douville et al., 1995), but this is not used in this study.

The ISBA scheme predicts the evolution of five prognostic variables: the surface temperature $T_s$, the mean surface temperature $T_2$, the superficial volumetric water content $w_g$, the mean volumetric water content $w_2$ and the water amount of an interception reservoir $W_r$, as follows:

$$\frac{\partial T_s}{\partial t} = C_T(R_n - H - LE) - \frac{2\pi}{\tau}(T_s - T_2), \tag{1}$$

$$\frac{\partial T_2}{\partial t} = \frac{1}{\tau}(T_s - T_2), \tag{2}$$

$$\frac{\partial w_g}{\partial t} = \frac{C_1}{\rho_w d_1}((1-\text{veg}) \times P - E_g) - \frac{C_2}{\tau}(w_g - w_{geq}) \qquad 0 \leqslant w_g \leqslant w_{sat}, \tag{3}$$

$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}((1-\text{veg}) \times P - E_g - E_{tr}) - \frac{C_3}{\tau}\max[0, (w_2 - w_{fc})] \qquad 0 \leqslant w_2 \leqslant w_{sat}, \tag{4}$$

$$\frac{\partial W_r}{\partial t} = \text{veg} \times P - E_r - R_r \qquad 0 \leqslant W_r \leqslant W_{r\max}, \tag{5}$$

where the variables $R_n$, $H$, and $LE$ represent the net radiative, sensible and latent fluxes respectively. $P$, $E$, and $R$ correspond to precipitation, evaporation (or transpiration) and runoff rates.

The subscripts g, tr, and r stand for "ground level", "transpiration", and "interception reservoir".

The mean characteristics of the scheme are presented below:

1. The surface is represented by a fraction of bare ground $(1-\text{veg})$ and a fraction of soil covered by vegetation (veg).

2. A single heat budget is considered for the surface and the vegetation. The coefficient $C_T$ represents the average heat capacity of bare ground and vegetation. The mean surface temperature $T_2$ is driven by the surface temperature $T_s$ with $\tau$ a restore time period of one day.

3. The soil water content is represented by 2 reservoirs: the superficial reservoir of thickness $d_1$ (1 cm) and the total soil reservoir of thickness $d_2$ (of order $1m$) containing the root-zone from which the vegetation extracts water. The total reservoir contains the superficial reservoir (Fig. 1). The volumetric water contents $w_g$ and $w_2$ are equivalent to the corresponding quantities of water $W_s$ and $W_p$ ($W_s = d_1 w_g$ and $W_p = d_2 w_2$). An interception reservoir is used to modelize the amount of water intercepted by the leaves. The coefficients $C_1$, $C_2$, $C_3$ describe the hydraulic properties of the soil affecting the infiltration at the surface, the subsurface conductivity and the gravitational drainage, respectively. The terms $C_2/\tau$ and $C_3/\tau$ are of order 1 and 10 days. The parameters $w_{\text{wilt}}$, $w_{\text{fc}}$, $w_{\text{sat}}$ are the volumetric moisture contents at wilting point (below which the plant transpiration stops), at field capacity (filled reservoir) and at saturation (maximum water content). $w_{\text{geq}}$ is the surface volumetric moisture at the balance of gravity and capillary forces. The parameters $w_{\text{wilt}}$, $w_{\text{fc}}$, $w_{\text{sat}}$ depend on soil texture, and $C_T$, $C_1$, $C_2$, $C_3$, $w_{\text{geq}}$ on soil texture and soil moisture.

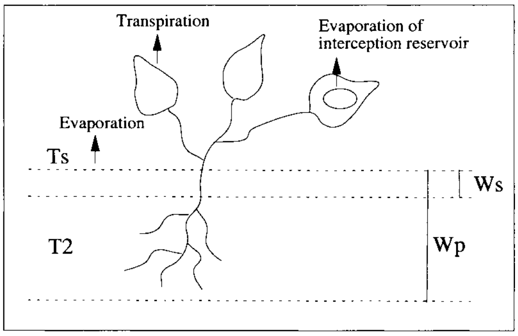


*Fig. 1.* Representation of the evaporation fluxes in the surface scheme ISBA.

The total evaporation is the sum of surface evaporation from bare ground ($E_g$), transpiration of vegetation ($E_{tr}$) extracting water from the total reservoir, evaporation of interception reservoir ($E_r$) (Fig. 1). The expressions of the evapotranspiration fluxes are:

$$E_g = (1-\text{veg})\rho_a C_H V_a (h_u q_{\text{sat}}(T_s) - q_a), \tag{6}$$

$$E_r = \rho_a \,\text{veg}\, \frac{\delta}{R_a} (q_{\text{sat}}(T_s) - q_a), \tag{7}$$

$$E_{tr} = \rho_a \,\text{veg}\, \frac{1-\delta}{R_a + R_s} (q_{\text{sat}}(T_s) - q_a), \tag{8}$$

where $\rho_a$, $V_a$, $q_a$ are respectively the air density, the wind speed and the specific humidity at the lowest atmospheric level. $C_H$ is the drag coefficient depending upon the thermal stability of the atmosphere and the surface roughness. $q_{\text{sat}}(T_s)$ is the saturated specific humidity at the temperature $T_s$. $\delta$ is the fraction of the foliage covered by the interception reservoir. $R_a = (C_H V_a)^{-1}$ is the aerodynamic resistance.

The interception reservoir has a very small capacity and contains water only a few hours after a rainfall. It affects surface fluxes during a very short time only. For that reason its water content is not analysed in this study.

The surface evaporation from bare ground ($E_g$) and the transpiration ($E_{tr}$) depend on the soil water contents $w_g$ and $w_2$ respectively through the relative humidity $h_u$ at the ground surface and the canopy surface resistance $R_s$ (Fig. 2).

## 3. The tangent linear model

The temporal evolution of a model can be written as: $\boldsymbol{x} = M[\boldsymbol{x}_0]$ with $\boldsymbol{x}_0$ the initial vector of control variables, $\boldsymbol{x}$ this vector at time $t$ and $M$ the integration of the model during the period $[0, t]$. The tangent linear model is the linearization of the model around a given state $\boldsymbol{x}_0$, it is represented by the Jacobian matrix $\boldsymbol{M}(\boldsymbol{x}_0)$ of the function $M$. A variation of the initial control variables will correspond a variation of the model prediction at time $t$ according to: $\delta\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{x}_0)\delta\boldsymbol{x}_0$. The tangent linear model can be used to analyse the impact of small disturbances. In the following, we will first test the validity of the tangent linear hypothesis.

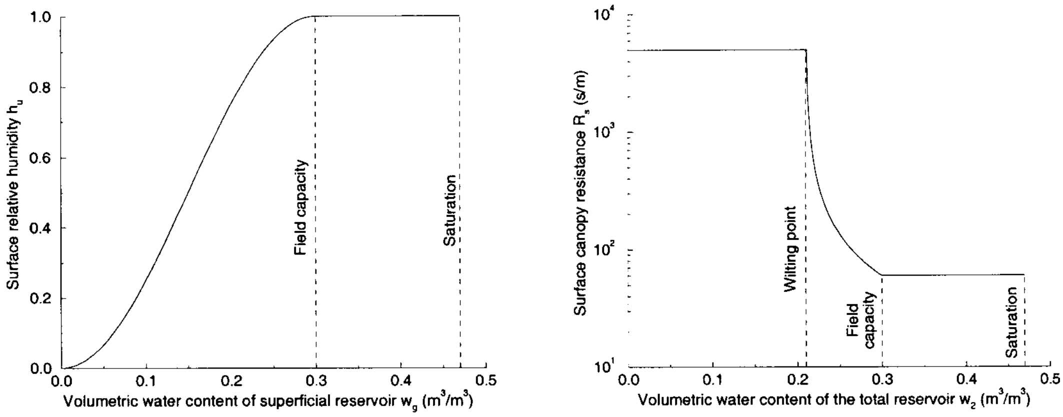


*Fig. 2.* General shapes of surface relative humidity $h_u$ as a function of $w_g$ (left) and surface canopy resistance $R_s$ as a function of $w_2$ (right).

Next the tangent linear model is used to evaluate the influence of surface parameters on surface fluxes and screen level parameters.

### 3.1. *Validation of the tangent linear hypothesis*

This 1D study is realized for a grid point of the 3D model located in south of France on 11 March 1997. This period is taken to have a solar incident radiative flux at the surface close to a mean value over the year (equinox). The situation is anticyclonic (low wind, no clouds). The initial soil variables are: $T_s = 277.5$ K, $T_2 = 283.3$ K, $W_s = 2.21$ mm, $W_p = 576$ mm. Table 1 displays the ISBA structure parameters for this experiment.

The tangent linear model describes the impact of an initial perturbation on the forecast. It corresponds well to the real forecast done with the perturbed state as long as the initial perturbation is small. The aim here is to evaluate if the tangent linear hypothesis is still valid for a mean perturbation. The magnitude of the chosen perturbation is of the same order as the root mean square (rms) forecast errors on soil variables: $\delta T_s = +2°$, $\delta T_2 = +2°$, $\delta W_s = +0.5$ mm and $\delta W_p = +100$ mm. These values are roughly estimated according to the errors of the land-surface scheme and of its forcings (precipitations, downward radiative fluxes, ...). In Fig. 4 the different curves $M[\boldsymbol{x}_0]$, $M[\boldsymbol{x}_0 + \delta\boldsymbol{x}_0]$ and $M[\boldsymbol{x}_0] + \boldsymbol{M}(\boldsymbol{x}_0)\delta\boldsymbol{x}_0$ are plotted for a 48 h forecast. $\boldsymbol{x}_0$ represents the initial surface state, $\delta\boldsymbol{x}_0$ the perturbation applied to all control variables, $M$ the integration of the model and $\boldsymbol{M}$

Table 1. *The soil and vegetation structure parameters of ISBA*

| | |
|---|---|
| depth of soil column ($d_2$) | 2.32 m |
| minimum leaf surface resistance | 65 s m$^{-1}$ |
| leaf area index | 1.55 m$^2$ m$^{-2}$ |
| vegetation fraction cover (veg) | 57% |
| roughness length for momentum | 0.86 m |
| roughness length for heat | 0.086 m |
| surface albedo | 18% |
| emissivity | 96% |
| wilting point volumetric moisture content ($w_{wilt}$) | 0.208 m$^3$ m$^{-3}$ |
| field capacity volumetric moisture content ($w_{fc}$) | 0.298 m$^3$ m$^{-3}$ |
| saturated volumetric moisture content ($w_{sat}$) | 0.467 m$^3$ m$^{-3}$ |
| quantity of water in the total reservoir at the wilting point | 482 mm |
| quantity of water in the total reservoir at the field capacity | 691 mm |
| quantity of water in the total reservoir at the saturation | 1080 mm |

the tangent linear model. The surface fluxes are defined positive downwards.

Ideally, $M[\boldsymbol{x}_0 + \delta\boldsymbol{x}_0]$ and $M[\boldsymbol{x}_0] + \boldsymbol{M}(\boldsymbol{x}_0)\delta\boldsymbol{x}_0$ should be similar. From Fig. 3 it can be seen that they are relatively close. Nevertheless, some discrepancies can be observed locally, more important on $2m$ relative humidity. The non-linearities of the model are responsible for these differences. An important one corresponds to the formulation of the exchange coefficients. The derivatives become very large when the Richardson number

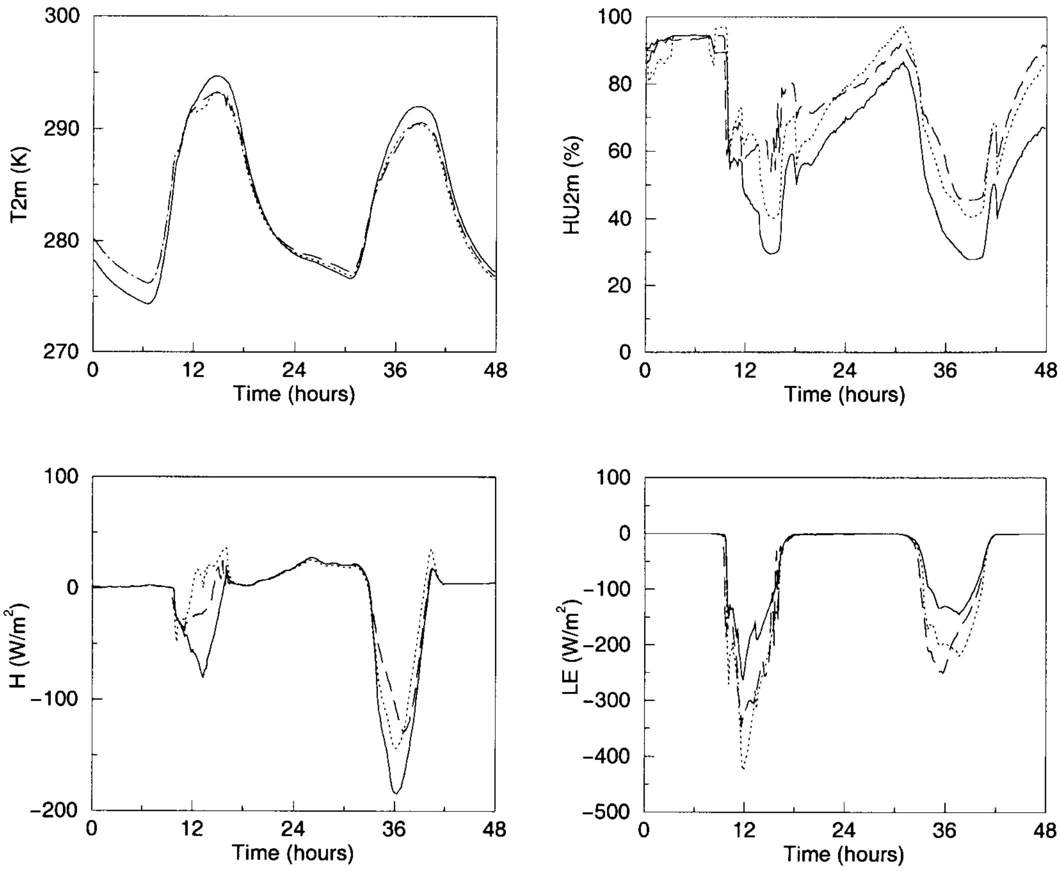


*Fig. 3.* Evolution of the screen level parameters (temperature $T_{2m}$ and relative humidity $HU_{2m}$ at 2*m*) and of the surface fluxes (sensible *H* and latent LE heat fluxes) during 48 h. The solid, the long-dashed and the dotted lines correspond respectively to $M[x_0]$, $M[x_0+\delta x_0]$ and $M[x_0]+\mathbf{M}(x_0)\delta x_0$.

approaches zero and create strong oscillations in the tangent linear model. This situation occurs twice a day when the stability of the atmosphere changes sign (Janisková et al., 1997). This problem is solved by reducing the derivatives when the Richardson number is close to zero. This modification does not degrade the quality of the tangent linear model elsewhere.

This experiment was repeated several times with different initial perturbations. Similar results were obtained showing that the tangent linear hypothesis is globally well satisfied up to 48 h. The validity of this hypothesis was not tested on a longer period because then the 1D model forecasts are less and less accurate.

### 3.2. *Sensitivity studies*

In this chapter, the influence of each surface parameter on the surface fluxes and the screen level parameters is evaluated with the tangent linear model on the same situation and the same geographical point as described above. In Fig. 4 the results of four experiments realized with an initial perturbation separately on $T_s$, $T_2$, $W_s$ or $W_p$ are represented.

We can observe the consequences of the surface perturbations on the surface fluxes. These fluxes are very sensitive to the soil water contents. An increase of the soil moisture increases the latent heat flux and reduces the sensible heat flux. The influence lasts longer than two days for the

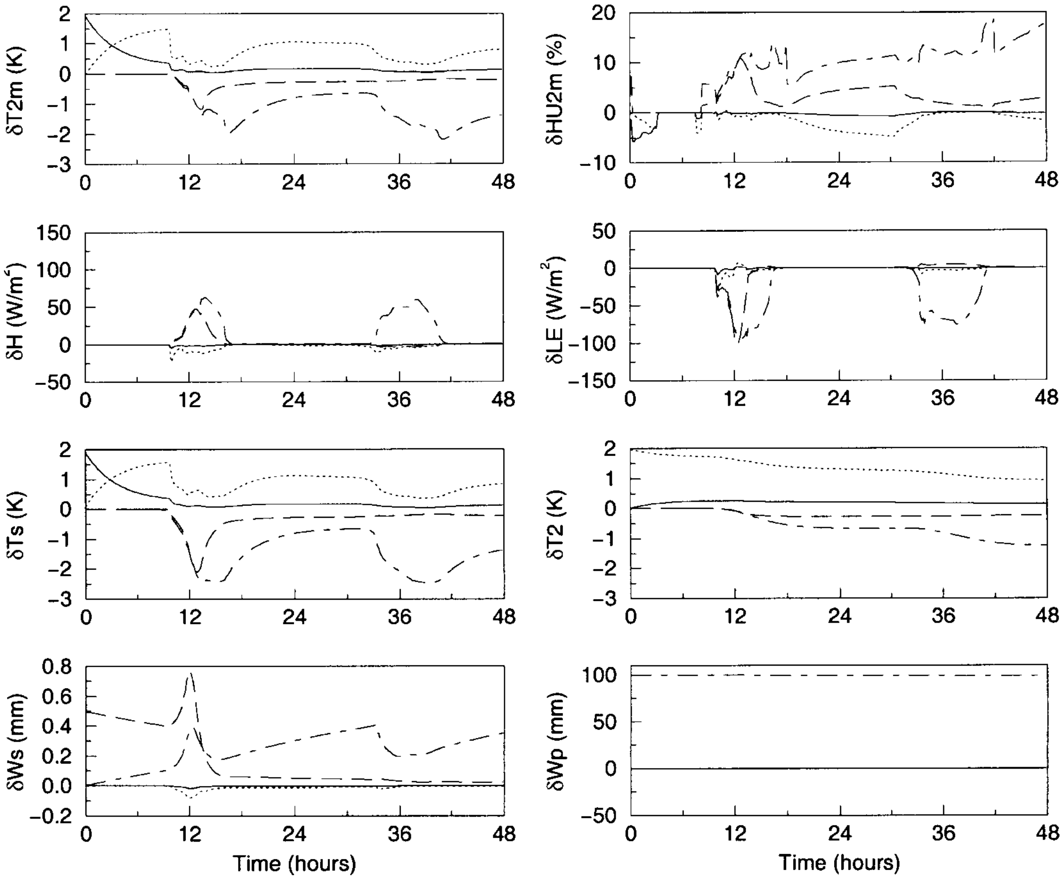


*Fig. 4.* Sensitivities studies performed with the tangent linear model. The solid, the dotted, the long dashed and the dot-dashed lines are related respectively to the following perturbations: $\delta T_s = +2°$, $\delta T_2 = +2°$, $\delta W_s = +0.5$ mm and $\delta W_p = +100$ mm.

perturbation on the total soil water content $W_p$. The impact of the perturbation on the superficial reservoir content is limited to the first day. Because of the small capacity of the superficial reservoir the evaporation dries this reservoir in a few hours under clear sky conditions. The influence of the mean soil temperature on surface fluxes is very low, whereas perturbing the surface temperature has no impact at all.

During the time of solar heating the screen level parameters are forced by the surface fluxes and therefore mainly depend on the soil water contents. During night with low wind the boundary layer is stable. Under these conditions the soil temperatures have an influence on the 2*m* temperature. The action of surface temperature is disappearing rapidly whereas the mean soil temperature has still an influence after two days. In our case the soil moisture reservoirs still have a small impact during night, because the air at 2*m* does not reach the saturation point. When it happens the 2*m* relative humidity becomes independent of the soil humidities.

The four last panels in Fig. 4 inform us about the time scale evolutions of the surface variables. The initial perturbations on the surface temperature and on the superficial water reservoir disappear rapidly. An initial bias on mean surface temperature is halved after two days. The initial error on the total soil water content is virtually unchanged after 48 h.

This short study shows that the soil water contents have an important influence on the surface fluxes and screen level parameters during daytime. Therefore it is important to analyse these two parameters, in particular the total moisture because of its much longer time scale evolution. Furthermore the superficial reservoir has a small capacity, an initial error is corrected after two days if the total water content is correct because of the subsurface conductivity. It seems also important to analyse the mean soil temperature because it has an influence on the $2m$ temperature during stable conditions and a quite long time scale. The influence of the surface temperature is small and short in time.

The results of this sensitivity study refering to a 48 h assimilation interval rely on the assumption that the 1D approach is valid for this time span. We can see in Fig. 4 that the perturbation on the total soil moisture has the same impact on the surface fluxes for the two days, but an impact steadily increased on the atmospheric parameters. This effect would be probably a bit less important in reality. The 1D model is indeed not able to relax to a balanced state because all remaining forces (radiation, advection, ...) are prescribed. For this reason the 1D model should be use only on a short period (one or two days maximum) and in anticyclonic meteorological conditions.

## 4. Variational environment

The variational method minimizes the cost function $J$ that measures the weighted sum of squares of distances $J^{\mathrm{b}}$ to the background (generally a short range forecast) and $J^{\mathrm{o}}$ to the observations distributed over the assimilation period. The control variable $\boldsymbol{x}$ contains the parameters to be analysed (here $T_{\mathrm{s}}$, $T_2$, $W_{\mathrm{s}}$, $W_{\mathrm{p}}$).

$$\begin{aligned} J(\boldsymbol{x}) &= J^{\mathrm{b}}(\boldsymbol{x}) + J^{\mathrm{o}}(\boldsymbol{x}) \\ &= \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}})^{\mathrm{T}}\boldsymbol{B}_0^{-1}(\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}) \\ &\quad + \frac{1}{2}\sum_{i=1}^{n}(H_i[M_i[\boldsymbol{x}]]-\boldsymbol{y}_i)^{\mathrm{T}}\boldsymbol{R}_i^{-1} \\ &\quad \times (H_i[M_i[\boldsymbol{x}]]-\boldsymbol{y}_i), \end{aligned} \tag{9}$$

with: $\boldsymbol{x}$ is the control variable of the model at time $t_0$; $\boldsymbol{x}^{\mathrm{b}}$ is a forecast of the model at time $t_0$ (background); $\boldsymbol{B}_0$ is the covariance matrix of background error; $\boldsymbol{y}_i$ is the observation vector at time $t_i$; $H_i$ is the observation operator at time $t_i$; $M_i$ is the integration of the model during the period $[t_0, t_i]$; $\boldsymbol{R}_i$ is the covariance matrix of observation error at time $t_i$.

The solution is sought iteratively by performing several evaluations of the cost function and of its gradient in order to approach the minimum using a quasi-Newton method. A pre-conditioning is applied in order to improve the convergence of the minimizer (Gilbert and Lemaréchal, 1989). An incremental approach is not necessary because of the simplicity of the model (Courtier et al., 1995). The computation of the gradient needs the adjoints of the model $\boldsymbol{M}_i^{\mathrm{T}}$ and of the observation operator $\boldsymbol{H}_i^{\mathrm{T}}$ defined as the adjoint of the Jacobian matrix $\boldsymbol{M}_{\mathrm{i}}$ and $\boldsymbol{H}_{\mathrm{i}}$, respectively.

$$\begin{aligned} \nabla J(\boldsymbol{x}) &= \boldsymbol{B}_0^{-1}(\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}) \\ &\quad + \sum_{i=1}^{n}\boldsymbol{M}_i^{\mathrm{T}}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{R}_i^{-1}(H_i[M_i[\boldsymbol{x}]]-\boldsymbol{y}_i), \end{aligned} \tag{10}$$

with: $\boldsymbol{H}_i^{\mathrm{T}}$ is the adjoint of the observation operator at time $t_i$; $\boldsymbol{M}_i^{\mathrm{T}}$ is the integration of the adjoint model during $[t_i, t_0]$.

Simple covariance matrices of background and observation errors are chosen. All the non-diagonal terms are equal to zero. For the matrix $\boldsymbol{B}_0$ the diagonal corresponds to the rms forecast errors on the surface parameters. For the matrix $\boldsymbol{R}_i$ it is the rms analysis errors on the $2m$ parameters. For an operational implementation it would be interesting to have a better evaluation of these matrices. This requires specific work not investigated here.

$$\boldsymbol{B}_0 = \begin{pmatrix} \sigma_{T_{\mathrm{s}}}^2 = 2^2 & 0 & 0 & 0 \\ 0 & \sigma_{T_2}^2 = 2^2 & 0 & 0 \\ 0 & 0 & \sigma_{W_{\mathrm{s}}}^2 = 0.5^2 & 0 \\ 0 & 0 & 0 & \sigma_{W_{\mathrm{p}}}^2 = 100^2 \end{pmatrix} \tag{11}$$

$$\boldsymbol{R}_i = \begin{pmatrix} \sigma_{T_{2m}}^2 = 1^2 & 0 \\ 0 & \sigma_{HU_{2m}}^2 = 0.1^2 \end{pmatrix} \tag{12}$$

The minimizer performs several tens of iterations to find the minimum of the cost function. Because of the computational cost the minimization is ended when the state variable is close enough to the minimum. To evaluate the optimal number of iterations many variational analyses were realized. According to the statistical results presented in Fig. 5 it was decided to limit the number of iterations to 10.

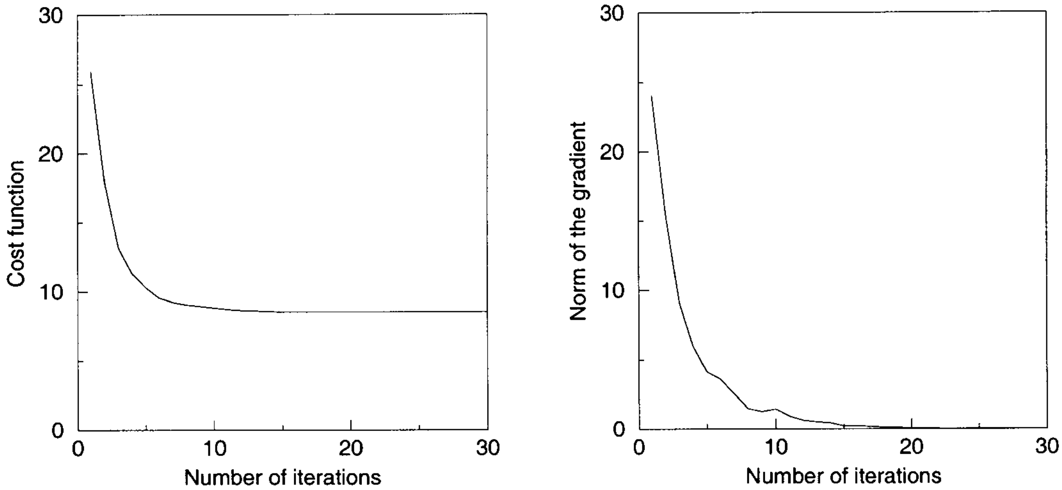


*Fig. 5.* Decrease of $J(x)$ and $\|\nabla J_x\|$ in relation with the number of iterations for 100 experiments.

## 5. General shape of the cost function

### *5.1. Topology of* $J^o$

The observations of temperature and relative humidity at $2m$ are generated by the 1D model and a gaussian noise with a rms of 1°C on $T_{2m}$ and 10% on $HU_{2m}$ is added to simulate observational errors. We are using observations every 3 h on a 24 h assimilation period. The initial state from which the observations are generated is called the true or real initial state $\boldsymbol{x}^t$.

In Fig. 6, different cross-sections of the cost function with $J^o$ only are presented. They contain all $\boldsymbol{x}^t$ and correspond to the possible couples between $T_s$, $T_2$, $W_s$, $W_p$. The soil temperature and soil moisture analyses are mainly uncoupled. The influence of soil temperatures on the cost function is approximately quadratic. The most important variations of the cost function are related to $W_s$ and $W_p$, showing again the sensitivity of $2m$ parameters to soil moisture. The cross-section corresponding to the plan ($W_s$, $W_p$) has an irregular shape as found in Callies et al. (1997). It is surrounded by flat zones corresponding to dry or saturated reservoirs. There is a valley of low values where the gradient is very small. This indicates an ambiguity for the repartition of water in the two reservoirs. As long as the values of ($W_s$, $W_p$) remain in the valley the response of the atmosphere at $2m$ is alike on the assimilation period. In this particular case there is a secondary minimum.

### *5.2. Convergence of the minimization*

In this subsection the cost function contains only the term of distance to the observations $J^o$. The aim is to evaluate whether the method is able to realize a good analysis, i.e., to approach closely the real initial $\boldsymbol{x}^t$ state even with noise on observations and a bad initial state. As seen before, the evapotranspiration is constant and the $2m$ parameters are not dependent any more on soil moistures when the soil reservoirs are dry (below $w_{wilt}$ for $W_p$) or above the field capacity. The gradient of $J^o$ is equal to zero over these zones and the minimizer stops its search. To solve that problem one solution could be to modify the parametrizations of the land-surface scheme to always keep the $2m$ parameters dependent on soil parameters or to add a quadratic term $J^c$ to the cost function representing the distance to the climatology.

Fig. 7 presents the results in the plan ($W_s$, $W_p$) of three assimilations with $J^o$ only starting from different initial conditions. The real initial state $\boldsymbol{x}^t$ is represented by a white cross. In the first case the minimizer does nothing because the gradient of the cost function is equal to zero. The analysis is close to the real initial state $\boldsymbol{x}^t$ when starting from the initial state 2. Case 3 corresponds to a bad analysis. The minimizer has found a secondary minimum. Several assimilations starting from different initial states were realized. The analysis were either close to $\boldsymbol{x}^t$ or to the secondary minimum. These experiments validate the variational

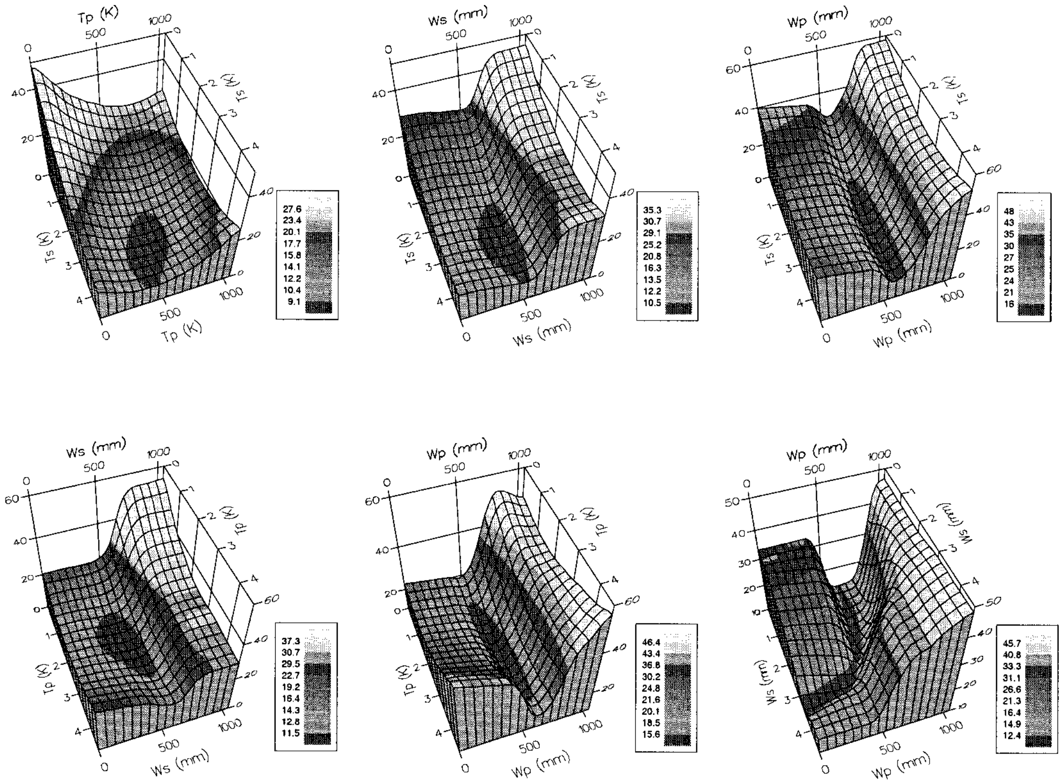

*Fig. 6.* Representation of several cross-sections of the cost function $J^o$ for simulated observations every 3 h and an assimilation period of 24 h.

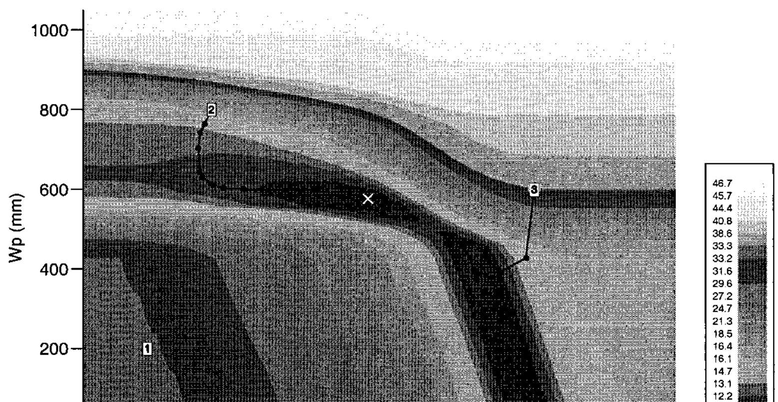


*Fig. 7.* Three variational assimilations with $J^o$ only starting from different initial conditions represented by the numbers. The cross corresponds to the real initial state.

environment (adjoint model, minimization, tangent linear hypothesis) and point out the fact that the cost function might not be convex.

In Fig. 8 the evolution of the model is plotted starting from the analysed states of cases 2 and 3. The screen level parameters and the surface heat fluxes are plotted. The total evaporation is splitted in transpiration and evaporation from bare ground. In both cases the observations are fitted quite well during the assimilation period. The sensible and latent heat fluxes are similar but the ratio between bare ground evaporation and transpiration are very different. In this particular case the latent heat flux can be adjusted with the evaporation from bare ground only, the transpiration being stopped. This ambiguity is a problem, because a realistic partitioning of water is necessary to have a good forecast during several days.

### 5.3. *First experiments with $J^o$ and $J^b$*

It is interesting to use the background term in the cost function for many reasons:

- The background term $J^b$ takes into account the information coming from the past. Until recently the surface analysis was so crude that the surface variables contained a lot of errors. With improved surface analysis and new land-surface schemes some pertinent information will be available in the previous forecast. This will be surely

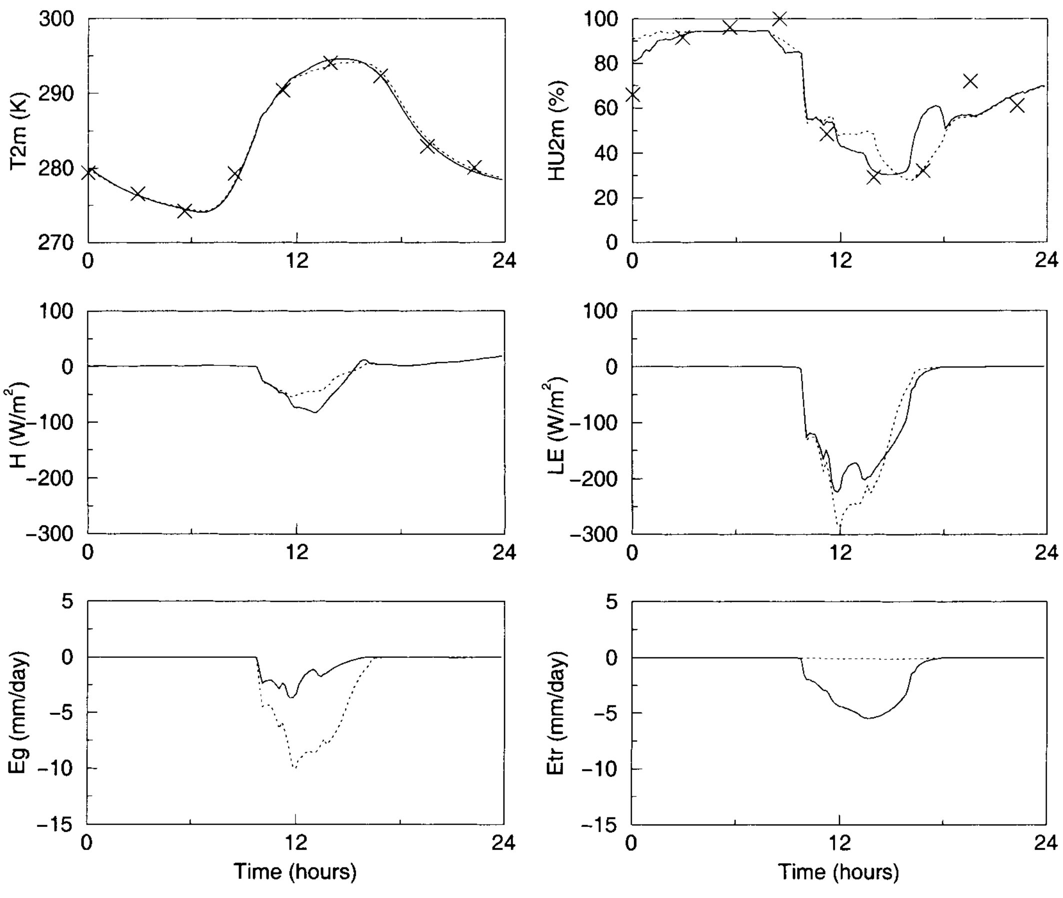


*Fig. 8.* Comparaison of the forecasts obtained from the analysis states of cases 2 (solid line) and 3 (dotted line). The crosses represent the simulated observations of $T_{2m}$ and $HU_{2m}$ every 3 h. $E_g$ and $E_{tr}$ stand respectively for the surface evaporation from bare ground and the transpiration.

the case for soil parameters which have a long time scale evolution period like $T_2$ or $W_p$. To improve the use of background information it is important to determine precisely the covariance matrix of background errors. The evaluation of this matrix $B_0$ is difficult because strongly dependent on the previous meteorological conditions. Mean values are chosen as described before.

- The variational analysis using $J^o$ only is dangerous because its result can be very bad in case of non informative observations.
- The shape of the cost function is more quadratic with $J^b$.

The use of a background term in the cost function raises the problem of the initialization of the control variable $\boldsymbol{x}^i$ before the minimization. Usually a previous forecast $\boldsymbol{x}^b$ is taken because it represents our best knowledge about the control variable. But in our case the gradient of $J^o$ with respect to the soil moistures in $\boldsymbol{x}^b$ may be equal to zero (in case of very dry or wet soils). The analysis will then have no effect because the initial state corresponds to a local minimum. To prevent this situation, we have decided to keep $\boldsymbol{x}^b$ as initial state when $0.5w_{\text{wilt}} \leqslant w_g \leqslant w_{\text{fc}}$ and $w_{\text{wilt}} \leqslant w_2 \leqslant w_{\text{fc}}$ (so most of the cases). When this condition is not satisfied medium values for the initial soil moistures are chosen ($w_g = w_2 = (w_{\text{wilt}} + w_{\text{fc}})/2$). To evaluate the improvement of the resetting 100 assimilation experiments ($J^b + J^o$) were realized starting from forecasts $\boldsymbol{x}^b$ different from the real initial state $\boldsymbol{x}^t$ (eq. (13)). The assimilation period is 24 h and the observations are simulated, with noise and present every 3 h. The statistics of the analyses are presented in Table 2. The analyses (1) and (2) represent respectively the results obtained without and with the resetting described before on the initial state $\boldsymbol{x}^i$. The improvement on $W_s$ and $W_p$ is important.

$$\boldsymbol{x}^b = \boldsymbol{x}^t + \begin{pmatrix} \varepsilon_G(0, \sigma^G_{T_s} = 2) \\ \varepsilon_G(0, \sigma^G_{T_2} = 2) \\ \varepsilon_G(0, \sigma^G_{w_g} = 0.5) \\ \varepsilon_G(0, \sigma^G_{w_2} = 100) \end{pmatrix} \tag{13}$$

where $\varepsilon_G(0, \sigma)$ is the realization of a gaussian variable with no bias and a rms error $\sigma$.

## 6. Quality of the analysis in different conditions

### 6.1. *Influence of the real soil water contents*

In this chapter the ability of the variational analysis to retrieve soil parameters for various soil water contents is evaluated. The same experiments as described before ($J^b + J^o$, simulated observations every 3 h with noise, assimilation period of 24 h) are performed with 5 different soil moistures: medium soil (A: $w_g = w_2 = (w_{\text{fc}} + w_{\text{wilt}})/2$), dry soil (B: $w_g = w_2 = w_{\text{wilt}} + 0.1(w_{\text{fc}} - w_{\text{wilt}})$), wet soil (C: $w_g = w_2 = w_{\text{wilt}} + 0.9(w_{\text{fc}} - w_{\text{wilt}})$), dry soil in surface and wet in depth (D), wet soil in surface and dry in depth (E). For each case 100 experiments were realized starting from different forecasts $\boldsymbol{x}^b$ (eq. (13)). The results are presented in Fig. 9.

The analysis for $T_s$ and $T_2$ is not sensitive to the initial water contents. Their rms errors are reduced by a factor 2 in comparison with the initial fields. It is interesting to note that the gain of the analysis on $W_s$ and $W_p$ is always present, but depends on the real soil water contents. The rms errors for $W_s$ and $W_p$ are correlated because of the ambiguity in the repartition of water in the two reservoirs. This ambiguity is more or less important depending on the real soil moistures.

Table 2. *Statistics of the analyses realized without (1) and with (2) resetting the soil moistures to medium values when the following condition is not satisfied* $0.5w_{wilt} \leqslant w_g \leqslant w_{fc}$ *and* $w_{wilt} \leqslant w_2 \leqslant w_{fc}$

| | | $\boldsymbol{x}^i$ | | Analysis (1) | | Analysis (2) | |
|---|---|---|---|---|---|---|---|
| | $\boldsymbol{x}^t$ | mean | rms error | mean | rms error | mean | rms error |
| $T_s$ | 277.5 | 277.6 | 1.89 | 277.5 | 0.77 | 277.5 | 0.76 |
| $T_2$ | 283.4 | 283.4 | 1.97 | 283.3 | 0.73 | 283.2 | 0.71 |
| $W_s$ | 2.21 | 2.29 | 0.52 | 2.28 | 0.46 | 2.20 | 0.32 |
| $W_p$ | 576.1 | 579.8 | 87.9 | 562.2 | 57.2 | 570.6 | 34.0 |

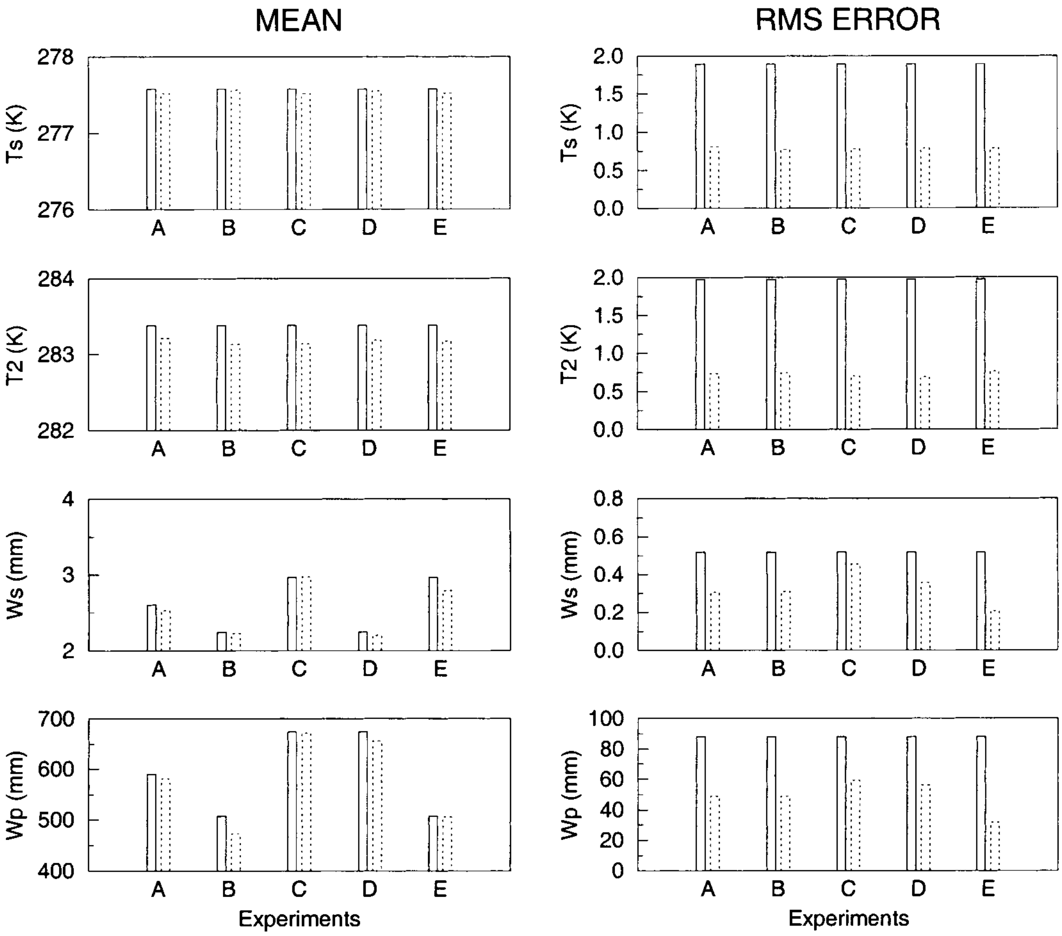


*Fig. 9.* Influence of the real soil water contents on the quality of the surface analysis. Five experiments are realized corresponding to: medium soil (A), dry soil (B), wet soil (C), dry soil in surface and wet in depth (D), wet soil in surface and dry in depth (E). The statistics (mean and rms error) of the initial (solid line) and analysed (dotted line) surface fields are plotted resulting from 100 cases for each experiment.

### *6.2. Influence of the vegetation*

The vegetation fraction cover (veg) is an important parameter for the soil water analysis. As a function of its value the superficial and the total reservoirs have very different effects on the global evaporation. On bare ground the latent flux depends mainly on the superficial water content, whereas in case of soil covered by vegetation it depends only on the mean soil water content.

The same kind of experiments as in the previous section are done with different fractions of vegetation: 0, 0.2, 0.5, 0.8, 0.99. The real soil water contents correspond to medium values. The results are presented in Fig. 10. The quality of the analysis for the surface temperatures is constant. On the other hand the influence of the vegetation fraction cover is important for the analysis of $W_s$ and $W_p$. When the parameter veg increases, the analysis of mean soil moisture gets better whereas the superficial soil moisture degrades. The analysis of soil water contents is good when their initializations are important. It is not a problem to have a bad superficial moisture with a soil covered by vegetation. The most important field ($W_p$) is always well analysed except in the case of bare ground.

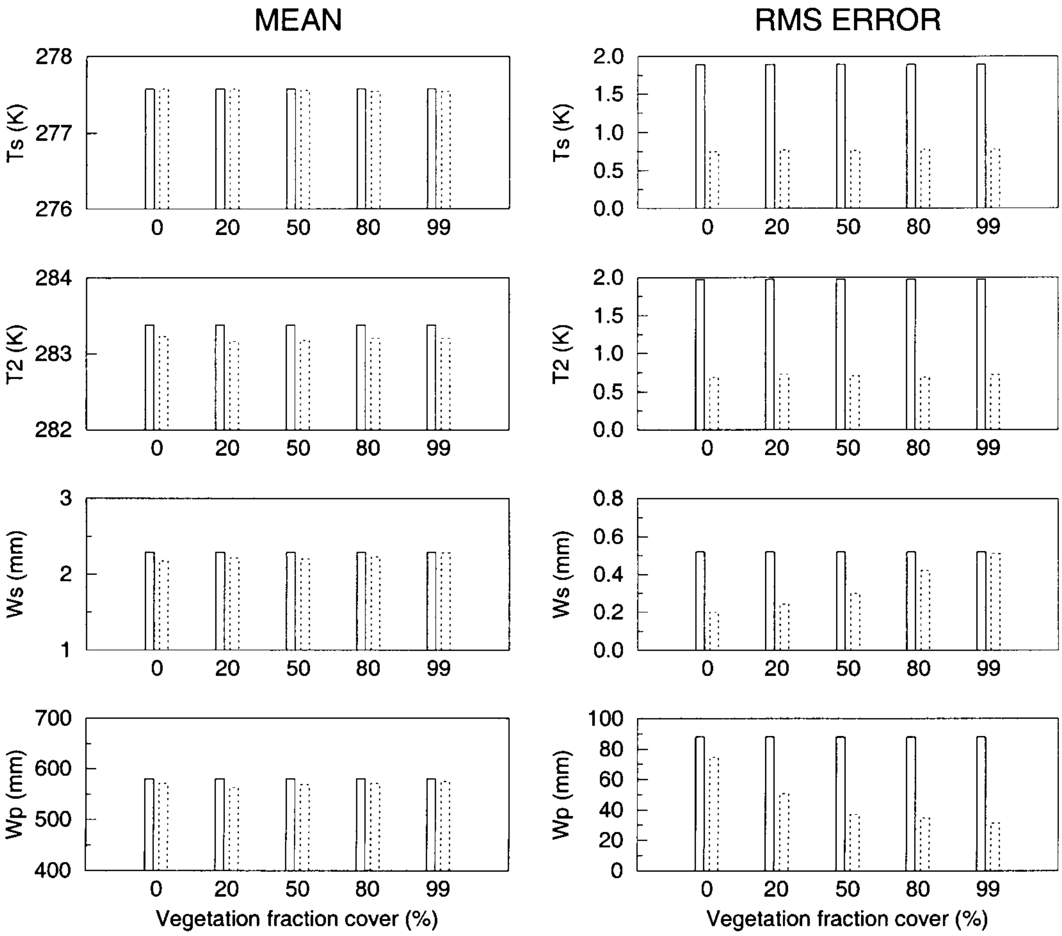


*Fig. 10.* Influence of the vegetation fraction cover on the quality of the surface analysis.

However even in that situation there is a small improvement of the mean water content through the subsurface conductivity. So an initial bias will be corrected after several assimilation days.

### 6.3. *Influence of the frequency of the observations and the length of the assimilation period*

In this subsection the influence of the frequency of the observations (1 h, 3 h, 6 h) and of the assimilation period length (24 h, 48 h) on the analysis is tested. The vegetation fraction cover and the initial soil water contents have medium values. The statistics of these analyses are presented in Fig. 11.

The consequences of the increase of the assimilation period depend mainly on the time scale evolution of the different surface parameters. The analysed fields are indeed comparable for the surface parameters ($T_s$ and $W_s$) and better for the mean parameters ($T_2$ and $W_p$). This improvement is particularly important for the mean soil water content because an assimilation period of 48 h reduces strongly the problem of ambiguity for the soil water repartition. According to the small capacity of the superficial reservoir the total evaporation the second day is mainly due to the mean soil water content. The analysis of $W_p$ is better with the pair 48 h/3 h (assimilation period/time frequency of observations) than with the pair 24 h/1 h despite a number of observations used less important. Ambiguities are better resolved when the atmospheric response to the respective

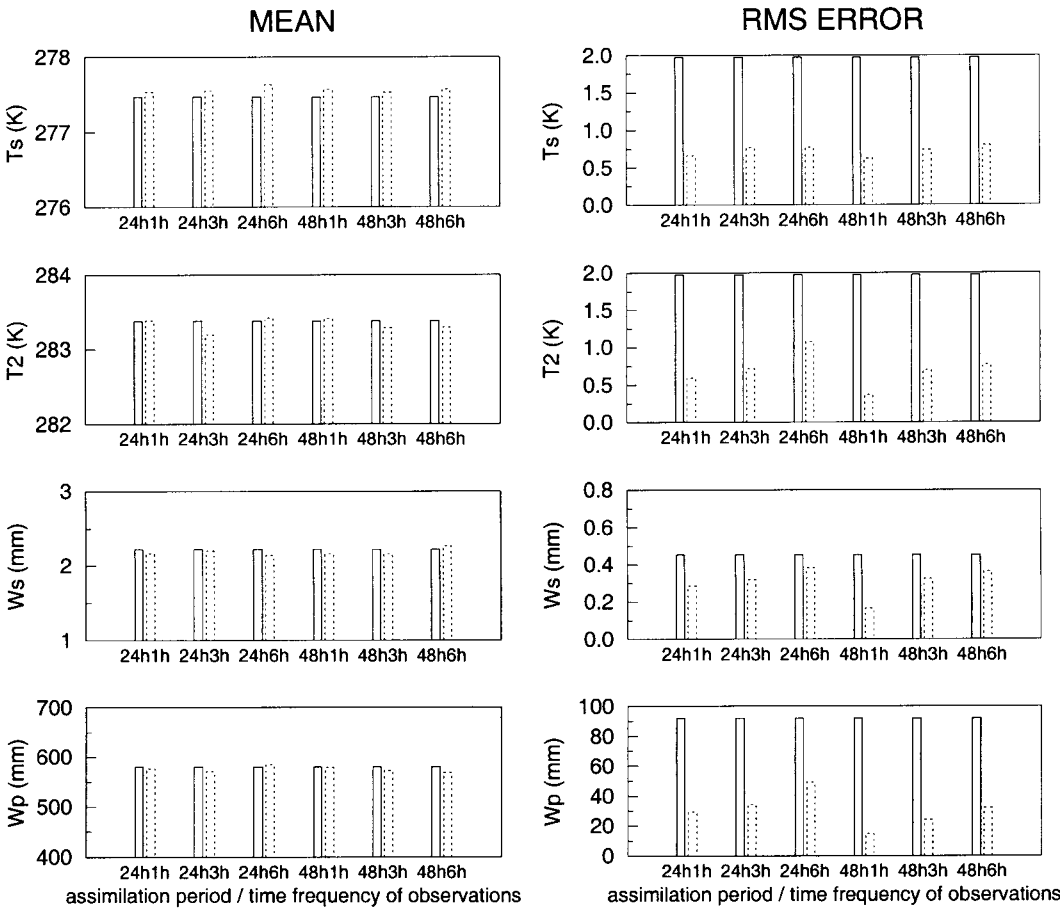


*Fig. 11.* Influence of the frequency of the observations (1 h, 3 h, 6 h) and the length of the assimilation period (24 h, 48 h) on the quality of the surface analysis.

soil parameters is different. The increase of the frequency of observations improves the quality of the analysis as planned, but it is important to note the strong improvement on the $W_p$ analysis when changing the frequency from 6 to 3 h.

## 7. Conclusions

A single-column model including vertical diffusion and land-surface schemes has been used. The other parametrizations (radiation, convection, precipitation, advection, ...) are prescribed from a 3D host model. This model is used to explore the impact of the surface parameters on screen level parameters and surface fluxes and to test the variational method for surface assimilation in the framework of identical twin experiments. All the results obtained here rely on the assumption of applicability of the 1D model during the assimilation interval. For this reason, we have chosen an anticyclonic situation (clear sky, low advection) and limited the assimilation period to 48 h.

First it is shown that the tangent linear hypothesis is valid for this simple model on a 48-h period. This is a prerequisite to generalize the results of the sensitivity studies for finite disturbances of the soil parameters. The tangent linear model is then used to evaluate the influence of soil variables on the surface fluxes and the $2m$ parameters. The mean soil water content appears to be the most important soil parameter to analyse for modelling the surface fluxes and the near surface parameters correctly. The superficial water content has a short time scale and is corrected by the subsurface conductivity if the total soil moisture is correct. The mean surface temperature is important too for a good forecast of the screen level parameters during stable conditions.

The characteristics of the variational method are then described. The analysis realizes the best compromise between the background information and the 2*m* observations according to the covariance matrix of corresponding errors. The general shape of the cost function with a 24-h assimilation period is presented. It reveals that the analysis of soil temperatures and soil moistures are mainly uncoupled in the sense of adjusting the 2*m* observations. The screen level parameters are indeed informative on the soil temperatures during the night and on the soil moistures during the day. The influence of the surface temperatures on the cost function is quadratic and so makes their initializations easier. This study points to the ambiguity in the repartition of water in the two soil reservoirs with a 24 h assimilation period. The latent heat flux can be adjusted in several ways with the evaporation from bare ground and the transpiration. This problem can be solved after several days by using the long time scale evolution of $W_p$ through the background term.

Finally, the efficiency of the variational assimilation is tested in different conditions with noisy observations. It was verified that this method is able to retrieve correctly the surface parameters with various soil water contents and vegetation fraction covers. It was shown that a frequency of 3 h in the observations improves the analysis of $W_p$ strongly in comparison to 6 h for an assimilation period of 24 h. An extension of the assimilation period to 48 h reduces the ambiguity in the repartition of water. Because of the small capacity of the superficial reservoir, the 2*m* observations of the 2nd day are informative on the mean soil water content only.

Some remaining problems have to be solved before an operational implementation of this surface analysis. First the 2*m* model errors independent of surface errors have to be taken into account. Otherwise the analysis will correct the distance to the observations but for wrong reasons and the forecasts will not be improved. It is for instance the case when the model makes an error on the cloudiness. To avoid this problem, it can be decided to realize the analysis in specific meteorological conditions (like anticyclonic) or to add some other fields in the control variable. Secondly it is important to have a better estimation of the covariance matrix of background errors. Adding this information will solve the problem of the ambiguity in the repartition of water. Thirdly the variational method should be tested on a long time period with real observations to study its stability.

## 8. Acknowledgements

Our thanks to E. Bazile and J. M. Piriou for their support with the one dimensional model and for many helpful discussions. We thank Dr. J. N. Thépaut for carefully reading the first draft of this paper and for providing valuable comments. We thank also Dr. J. F. Mahfouf from ECMWF for his constant support and advises. The final manuscript was significantly improved by the comments of two anonymous reviewers.

REFERENCES

Blackadar, A. K. 1976. Modeling the nocturnal boundary layer. Proc. 3rd Symp. Atmos. turbulence, diffusion and air quality, Boston, *Am. Meteorol. Soc*., 46–49.

Callies, U., Rhodin, A. and Eppel, D. P. 1998. A case study on variational soil moisture analysis from atmospheric observations. *J. of Hydrology*, 212–213. BAHC Special Issue, 95–108.

Courtier, P., Thépaut, J. N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367–1388.

Deardorff, J. W. 1977. A parametrization of the ground surface moisture content for use in atmospheric prediction models. *J. Appl. Meteor*. **16**, 1182–1185.

Douville, H., Royer, J. F. and Mahfouf, J. F. 1995. A new snow parameterization for the Météo-France climate model. Part I: validation in stand-alone experiments. *Climate Dynamics* **12**, 21–35.

Geleyn, J. F. 1988. Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus* **40A**, 347–351.

Geleyn, J. F., Bazile, E., Bougeault, P., Déqué, M., Ivanovici, V., Joly, A., Labbé, L., Piédelièvre, J. P., Piriou, J. M. and Royer, J. F. 1994. Atmospheric parametrization schemes in Météo-France's ARPEGE N.W.P. *Proceedings of the 1994 ECMWF Seminar on Physical parametrizations in numerical models,* 385–402, Reading.

Giard, D. and Bazile, E. 1997. Soil moisture assimilation in a global variable resolution NWP model. 13th Conference on Hydrology. *Am. Meteorol. Soc*., 163–166, Longbeach, California, 2–7 February.

Gilbert, J. C. and Lemaréchal, C. 1989. Some numerical experiments with variable storage quasi-Newton algorithms. *Mathematical Programming* **45**, 407–435.

Janisková, M., Thépaut, J. N. and Geleyn, J. F. 1999. Simplified and regular physical parametrizations for incremental four-dimensional incremental variational assimilation. *Mon. Wea. Rev*. 127, 26–45.

Jones, C. D. and Macpherson, B. 1995. *A hydrology corrections scheme for the mesoscale model using observed precipitation rates*. Forecasting Research Tech. Rep. 151, Meteorological Office, Bracknell, UK.

Louis, J. F. 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Meteorology* **17**, 187–202.

Louis, J. F., Tiedtke, M. and Geleyn, J. F. 1981. A short history of the operational PBL parameterization at ECMWF. *Workshop on the Planetary boundary layer parameterization*. ECMWF, 59–80, Reading, 25–27 November.

Macpherson, B. 1996. Initialisation of soil moisture in the operational mesoscale model. Forecasting Research Tech. Rep. 183, Meteorological Office, Bracknell, UK.

Mahfouf, J. F. 1991. Analysis of soil moisture from near-surface parameters: A feasibility study. *J. Appl. Meteor*. **30**, 1534–1547.

Mahfouf, J. F. and Noilhan, J. 1991. Comparative study of various formulations of evaporation from bare soil using *in situ* data. *J. Appl. Meteor*. **30**, 1354–1365.

Mahfouf, J. F. and Noilhan, J. 1996. Inclusion of gravitational drainage in a land surface scheme based on the force restore method. *J. Appl. Meteor*. **35**, 987–992.

Navascues, B. 1997. *Analysis of 2-m temperature and relative humidity*. HIRLAM Tech. Rep. 28, Norrköping, Sweden.

Noilhan, J. and Planton, S. 1989. A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev*. **117**, 536–549.

Paloscia, S. 1996. *Remote sensing of soil moisture. Radiation and water in the climate system*. T. J. E. Raschke (ed.). Springer, pp. 361–400.