Title: Deriving the surface soil heat flux from observed soil temperature and soil heat flux profiles using a variational data-assimilation approach.

Authors: R.J. Ronda, and F.C. Bosveld

Affiliation: Koninklijk Nederlands Meteorologisch Instituut (KNMI), De Bilt, Netherlands

Date of submission: 31st July 2008

Corresponding author:

Fred C. Bosveld

Koninklijk Nederlands Meteorologisch Instituut

Wilhelminalaan 10

NL-3732 GK De Bilt

Netherlands

Tel: +31302206788

Fax: +31302210407

E-mail: bosveld@knmi.nl

Tel: +31302206787

Abstract

A novel approach to infer surface soil heat fluxes from measured profiles of soil temperature, soil heat flux, and observations of the vegetation canopy temperature and the incoming short wave radiation, is evaluated for the Cabauw measurement facility in the Netherlands. The approach is a variational data-assimilation approach which uses the applied measurements to optimize on a daily basis parameter values of a model that describes the heat transport between the vegetation canopy and the surface and within the soil column. Inserting error characteristics that are either inferred from the field data itself or derived from literature, leads to valid estimates of the cost function for about 100 days in the year 2003. It appears that the approach gives values of the model parameters that compare well to values derived from literature, though at the end of 2003 values for the soil conductivity and the volumetric heat capacity of the soil start to differ from the literature values, possibly to specific soil characteristics and the extreme dryness of the summer of 2003. It appears that the model gives estimates of the surface soil heat flux that compare well with estimates using the currently operational Lambda approach, provided that the latter is adapted to account for the disturbance of the soil heat flux at the locations of the heat flux plates. Only when the surface soil heat flux is very small or very large, the new approach gives estimates of the surface soil heat flux that differ from the Lambda-approach.

1. Introduction

The soil heat flux quantifies transport of heat through planes perpendicular to the soil surface. At the surface it quantifies the energy transport (positive downward) through the interface between the air-vegetation continuum and the soil. The thus defined surface soil heat flux (G_0) is in balance with the net radiation (R_N) and the atmospheric transport (positive upward) of sensible (H) and latent heat (LE) (see Stull 1988).

During daytime and for vegetated surfaces G_0 is relatively small, about 10% of R_N (De Bruin and Holtslag 1982; Stull 1988). For sparsely vegetated and non-vegetated surfaces in semi-arid regions G_0 becomes a more important component of the surface energy budget (SEB) (Passerat de Silans et al. 1997; Heusinkveld et al. 2004). Also, during night time G_0 is important, because H and LE are in nocturnal conditions much smaller than during daytime. Consequently G_0 plays a significant role in the development of stable boundary layers (Van de Wiel et al. 2002) and in fog (Duynkerke 1991) and dew formation (Garrat and Segal 1988).

Local measurements of the soil heat flux at certain depths are obtained using heat flux plates (Portman 1958). To determine G_0 heat flux plates should be placed very close to the surface. This approach is however only applicable when a sensor is used which thermal properties are comparable to the thermal properties of the soil (Mayocchi and Bristow 1995). Using such a configuration, Heusinkveld et al. (2004) were able to measure the soil heat flux in the Negev desert.

In general, values of the thermal properties of the soil differ from thermal properties of heat flux plates that are used at measurement sites. G_0 is then inferred from a combination of heat flux plates that are placed at depths that are deeper than a few centimeters, and one or more soil temperature sensors. The estimation of G_0 involves an extrapolation towards the surface for which three methods are available:

1. the calorimetric method (Fuchs and Tanner 1968; Massman 1992; Heusinkveld et al. 2004), in which G_0 is calculated by integrating the one-dimensional heat conduction equation between the surface and the installation depth of the heat flux plate;
2. the harmonic analysis method (Van Wijk and de Vries 1963; Massman 1992; Heusinkveld et al. 2004), in which the G_0 follows from extrapolated harmonic

functions which are fitted to the vertical profile of measured soil heat fluxes and soil temperatures;

3. the Lambda method (De Bruin and Holtslag 1982), in which G_0 is calculated as the product of the observed surface gradient of soil temperature and an apparent thermal soil conductivity at the surface, which is derived from harmonic analysis of the observed soil heat fluxes and the observed surface temperature gradient.

Major disadvantage of the calorimetric method is that this method requires detailed knowledge of the thermal properties of the soil. Most notably the volumetric heat capacity of the soil, but also the soil heat conductivity of the soil as this is an important parameter for relating the heat flux over the plate to an undistorted soil heat flux (Philip 1961). In contrast, the harmonic analysis method does provide estimates of the thermal properties of the soil, provided that at least two measurements of the soil temperature are taken in conjunction with at least one measurement of the soil heat flux. Major disadvantage of this method is that it requires the slowest Fourier mode to be identifiable and measureable. Furthermore, it cannot calculate temporal variations in the soil characteristic that are more rapid than the identified slowest cycle. Major disadvantage of the Lambda method is that it determines the gradient at the surface by means of the difference between a measurement at the surface and a measurement at a few centimeters below the surface. In some situations, this might lead to underestimations or overestimations of the gradient at the surface.

In this paper, we pursue the use of an explicit model for the diffusion of heat from the vegetation canopy to the soil surface and within the soil column in conjunction with a variational data-assimilation technique. Using variational data assimilation techniques has a long history in weather forecasting (Daly 1991). Recently their application has been extended to 1D column models of the atmosphere (Lopez et al. 2006; Margulis and Entekhabi 2001) and land surface models (Boni et al. 2001). Main advantage of this method is that it provides optimal estimates of the soil characteristics in an user-defined data-assimilation window. Other advantages over the traditional methods to estimate the surface soil heat flux are that:

- the system provides a more flexible framework in which it is easy to incorporate different type of observations that contain information on the thermodynamic state of the soil in a consistent way;

- the system is able to incorporate instrumental corrections that depend on the thermal parameters of the soil;
- an objective quality assessment of the system and different sensors that are used within the system is possible;
- the system can be extended to incorporate known inhomogeneities of the soil.

The novel method is evaluated for measurements taken during the year 2003 at the Cabauw measurement facility in the Netherlands (Van Ulden and Wieringa 1996).

In section 2 the variational data-assimilation framework is described. Section 3 deals with the measurements that are used in the variational data-assimilation framework to estimate the surface soil heat flux. In section 4 the numerical setup for the validation study is given. In section 5, results of the validation study are presented and discussed. Finally, section 6 summarizes and concludes.

2. Variational data assimilation method

a. Forward model

The forward model consists of model equations for the system that is depicted schematically in Fig. 1. Core of the system is the soil column, which thermodynamic state is characterized by the vertically and temporally varying soil temperature, denoted by T. The soil temperature is found by integrating the one-dimensional diffusion (Fourier) equation from a starting time $t = t_0$ at which an initial profile for T is specified:

$$\frac{\partial T(z,t)}{\partial t} = \frac{\lambda}{c_p}\frac{\partial^2 T(z,t)}{\partial z^2}; \tag{1a}$$

$$T(z,t_0) = f_1(z), \tag{1b}$$

where c_p is the volumetric heat capacity of the soil, λ is the soil heat conductivity, z is the depth below the surface, and $f_1(z)$ gives the initial profile of T. Applying Eq. (1a) requires specification of boundary conditions at the top and bottom of the soil column. The upper boundary condition at the top of the soil column is formed by the surface soil heat flux. The surface soil heat flux follows from the transport of heat between the vegetation canopy and the surface of the soil, the long wave radiative flux difference between the vegetation canopy

and the soil surface, and the incoming short wave radiation that is absorbed at the soil surface. It is parameterized according to the formulation given by Duynkerke (1992), which translates to a Robin boundary condition for Eq. (1a) that reads as:

$$\left(\frac{\partial T(z,t)}{\partial z}\right)_{z=0} = -\frac{G_0(t)}{\lambda} = -\frac{1}{\lambda}\left(\Lambda\left(T_v(t) - \left(T(0,t)\right)\right) + \tau K_{in}(t)\right), \tag{1c}$$

where Λ is the skin conductivity, T_v is the vegetation canopy temperature, τ is the transmissivity of the vegetation, and K_{in} is the incoming short wave radiation flux. At the lower boundary, a Dirichlet boundary condition is implemented:

$$T\left(z_b, t\right) = f_2(t), \tag{1d}$$

where z_b denotes the depth at which the bottom of the soil column is located, and $f_2(t)$ specifies the soil temperature at the bottom of the column.

b. Optimization of parameters

To solve the system described by Eqs. (1a) to (1d) values for the parameters λ, c_p, Λ, and τ need to be specified. These are found using a variational data assimilation framework. Within this framework, we seek statistically optimal values for these parameters as a function of misfits between uncertain measurements and simulated equivalents, and uncertain prior parameter value estimates. Assuming that both measurements and prior value estimates can be considered as realizations drawn from a Gaussian distribution, deriving statistically optimal values for the parameters is equal to finding parameter values that minimize a cost function that reads:

$$J = (\mathbf{p} - \mathbf{p}_{\mathbf{p}r})^T C_p^{-1} (\mathbf{p} - \mathbf{p}_{\mathbf{p}r}) + \left(\mathbf{Z} - \mathbf{M}(\mathbf{p}, T(z,t;\mathbf{p})\right)^T C_v^{-1} \left(\mathbf{Z} - \mathbf{M}(\mathbf{p}, T(z,t;\mathbf{p})\right), \tag{2}$$

where $\mathbf{p}$ is the vector containing the parameter values: $\mathbf{p} = \left(\lambda, c_p, \Lambda, \tau\right)^T$; $\mathbf{p}_{pr}$ is the vector that contains the prior estimates of the parameter values: $\mathbf{p}_{pr} = \left(\lambda_{pr}, c_{p,pr}, \Lambda_{pr}, \tau_{pr}\right)^T$, where λ_{pr} is the prior value of λ, $c_{p,pr}$ is the prior value of c_p, Λ_{pr} is the prior value of Λ, and τ_{pr} is the prior value of τ; C_p is the error-covariance matrix of the prior values; $\mathbf{Z}$ is the vector containing the measurements; $\mathbf{M}$ is the vector that gives the modeled equivalents of the measurements that results from propagating the forward model, and C_v is the error-covariance matrix of the measurements.

To add the forward model as a hard constraint, its equations are adjoined to the cost function, forming the Lagrange function (Margulis and Entekhabi 2001,2004):

$$L = J + \int_{t_0}^{t_f} \int_{z_b}^{0} \mu^*(z,t) \left[\frac{dT(z,t)}{dt} - \frac{\lambda}{c_p} \frac{\partial^2 T(z,t)}{\partial z^2} \right] dz dt \tag{3}$$

where L is the Lagrange function, t_f is the end time of the integration (end time of data-assimilation window), and $\mu^*(z,t)$ is the Lagrange multiplier function.

Under the constraint of the forward model, minimizing the cost function is mathematically equivalent to minimizing the Lagrange function. The minimal value is obtained by setting the derivatives of the Lagrange function with respect to the state variables, the Lagrange multiplier function and the model parameters to zero:

$$\left(\partial L / \partial T(z,t)\right) = 0, \left(\partial L / \partial \mu^*(z,t)\right) = 0, \text{ and } \left(\partial L / \partial \mathbf{p}\right)^T = \mathbf{0}. \tag{4}$$

Differentiating the Lagrange function with respect to the Lagrange multiplier function, yields the forward model as given by Eqs. (1a) to (1d). Differentiating the Lagrange function with respect to the soil temperature field yields the adjoint model, which is given by:

$$\frac{\partial \mu^*(z,t)}{\partial t} = -\frac{\lambda}{c_p} \frac{\partial^2 \mu^*(z,t)}{\partial z^2} + \left(\frac{\partial J}{\partial T(z,t)} \right); \tag{5a}$$

$$\mu^*(z,t_f) = 0; \tag{5b}$$

$$\left(\frac{\partial \mu^*}{\partial z} \right)_{z=0} - \mu^*(0,t) \frac{\Lambda}{\lambda} = 0; \tag{5c}$$

$$\mu^*(z_b,t) = 0. \tag{5d}$$

The adjoint model is a terminal value problem. For given values for the skin conductivity, the soil heat conductivity, the volumetric heat capacity of the soil, and the transmissivity of the vegetation, the adjoint variables are found by integrating Eq. (5a) backward in time, using Eq. (5b) as terminal condition, and Eqs. (5c) and (5d) as boundary conditions.

Finally, differentiating the Lagrange function with respect to the model parameters yields:

$$\frac{\partial L}{\partial \lambda} = \frac{\partial J}{\partial \lambda} - \int_{t_0}^{t_f} \int_{z_b}^{z_t} \frac{T(z,t)}{c_p} \frac{\partial^2 \mu^*(z,t)}{\partial z^2} dz dt - \int_{t_0}^{t_f} \frac{f_2(t)}{c_p} \left(\frac{\partial \mu^*(z,t)}{\partial z} \right)_{z=z_b} dt \tag{6a}$$

$$\frac{\partial L}{\partial c_p} = \frac{\partial J}{\partial c_p} + \int_{t_0}^{t_f} \left(\int_{z_b}^{z_t} \frac{\lambda T(z,t)}{(c_p)^2} \frac{\partial^2 \mu^*(z,t)}{\partial z^2} \, dz - \frac{\mu^*(0,t)(\Lambda T_v(t) + \tau K_{in}) - \lambda f_2(t) \left(\frac{\partial \mu^*}{\partial z} \right)_{z=z_b}}{(c_p)^2} \right) dt \qquad (6b)$$

$$\frac{\partial L}{\partial \Lambda} = \frac{\partial J}{\partial \Lambda} + \int_{t_0}^{t_f} \frac{\mu^*(0,t) T_v(t)}{c_p} \, dt \qquad (6c)$$

$$\frac{\partial L}{\partial \tau} = \frac{\partial J}{\partial \tau} + \int_{t_0}^{t_f} \frac{\mu^*(0,t) K_{in}}{c_p} \, dt \qquad (6d)$$

Eqs. 1a to 1d, 5a to 5d, and 6a to 6d form a closed set of differential equations that can be solved. This yield values for soil temperature fields, Lagrange multiplier fields and model parameters that minimize the Lagrange function.

c. Numerical implementation

Obtaining analytical solutions for both the forward model and the adjoint model is not straightforward. The partial differential equation describing the forward model is homogeneous, but it includes inhomogeneous initial and boundary conditions. The adjoint model has homogeneous terminal and boundary conditions, but the partial differential equation describing its evolution includes non-linear source terms. Because in addition Eqs. (6a) to (6d) depend non-linearly on the model parameters, a numerical approach has been adopted to find optimal parameters.

First, the forward model is solved using a Crank-Nicolson finite difference scheme. This leads to a vector propagation equation that calculates the temperature for a set of discrete depths and time steps starting from a discretisized initial condition. Second, a discrete formulation of the Lagrange function is derived, whereby instead of the continuous forward model, the discretisized forward model is added as a constraint. Third, the discrete adjoint model is derived by differentiating the discrete formulation of the Lagrange function to the different components of the discrete temperature vector. Last, the discrete formulation for the Lagrange function is differentiated with respect to the model variables, yielding discrete equivalents of the Eqs. (6a) to (6d). Our approach can be identified as the adjoint of finite difference approach, as outlined by Sirkes and Tziperman (1997). The resulting equations are given in the appendix. Eqs. (A.1a) and (A.1b) describe the finite difference form of the

forward model, Eqs. (A.3a) and (A.3b) describe the resulting discrete adjoint model, while Eq. (A.4) summarizes the derived finite difference equivalents of Eqs. (6a) to (6d).

Because Eqs. (6a) to (6d) and their numerical equivalents represented by Eq. (A.4) are non-linear in the model parameters, finding the minimal values for the model parameters requires an iterative procedure. We follow Margulis and Entekhabi (2004) and employ a steepest descent technique to minimize L with respect to the control variables. This method is summarized in Algorithm 1, in which q refers to the iteration step counter , γ_q is the step size matrix, ε is the scalar convergence factor, while the subscript refers to the iteration step under scope.

Algorithm 1:

1 set q=1;

2. set $\mathbf{p}_q = \mathbf{p}_{pr}$;

3. calculate $\mathbf{T}_q(t_i), i = 1,2,3,..N$ using Eqs. (A.1a) and (A.1b);

4. calculate $\boldsymbol{\mu}_q(t_i), i = 1,2,3,..N$ using Eqs. (A.3a) and (A.3b);

5. calculate $\left\| \frac{\partial L}{\partial \mathbf{p}} \right\|_q$ using Eq. (Eq. A.4);

6. calculate $\mathbf{p}_{q+1} = \mathbf{p}_q - \gamma_q \left(\left(\frac{\partial L}{\partial \mathbf{p}} \right)_{\mathbf{y}=\mathbf{y}_q)} \right)^T$ using Eq. (A.4)

7. set q=q+1

8. repeat steps 3 to 7 until $\frac{\left\| \frac{\partial L}{\partial \mathbf{p}} \right\|_q}{\left\| \frac{\partial L}{\partial \mathbf{p}} \right\|_1} < \varepsilon$

3. Observations

a. General characteristics

The tower at the Cabauw measurement facility site is located at $51^\circ\ 58'$, $4^\circ 56'$, in the central Netherlands. It consists of flat grass meadows and ditches, surrounded by a flat area with villages, orchards and lines of trees (Beljaars and Bosveld 1997).

Jager et al. (1976) determined the soil in the Cabauw area using laboratory analysis of soil samples and inspection of a soil column in a 120-cm deep pit (see Beljaars and Bosveld 1997). The soil in Cabauw consists, from the surface downwards, of a 2 cm turf layer, a 18 cm upper clay layer that is relatively rich in organic matter, a 42 cm (from 18 cm to 60 cm) thick lower clay layer that is relatively poor in organic matter, a 15 cm thick mixed clay/peat layer, while underneath the soil consists of peat. Based on these characteristics the upper 18 cm of the upper layer is identified as B11 (fairly heavy clay) in the soil type classification for the Netherlands as proposed by Wösten et al. (1994).

b. Soil heat flux

In Cabauw, observations of the soil heat flux have been obtained at a so-called soil-terrain, located at about 100 m South of the main tower. Following Philip (1961), the soil heat flux follows from the heat flux that passes through a horizontally placed circular plate according to:

$$G = G_p\left(1 - 1.92\frac{T_p}{d_p}\left(1 - \frac{\lambda}{\lambda_p}\right)\right), \tag{7}$$

where G is the soil heat flux, G_p is the heat flux through the plate, T_p is the thickness of the plate, d_p is the diameter of the plate, and λ_p is the heat conductivity of the plate material. The heat flux through the plate is calculated as the product of the conductivity of the plate and the temperature gradient over the body of the plate. The latter is determined by the ratio of the temperature difference over the plate, which is measured using a thermopile, and the thickness of the plate.

At the Cabauw site six WS31S soil heat flux plates, manufactured by TNO-Delft are installed. The body of these plates consists of ceramic-plastic material, has a diameter of 110 mm, while it is 5 mm thick. These plates are buried at depths of approximately 5 and 10 cm below the surface, located at the three vertices of an equilateral triangle. First, for each depth and at each vertex 10-minute averages of the soil heat flux are obtained. Second, a linear average of the fluxes at the three vertices is performed so that average heat fluxes over the soil-terrain at the Cabauw site at depths of approximately 5 cm and 10 cm are obtained.

c. Soil temperature

Ten-minute averaged profiles of soil temperatures are measured at the centre of the equilateral triangle used to determine the average soil heat flux profile. The thermometers are KNMI home-made and consist of 35 cm long Nickel needle with a temperature-dependent electrical resistance around 500 Ohm. The thermometers were placed at depths of approximately 0.004 m, 0.02 m, 0.04 m, 0.08 m, 0.12 m, 0.20 m, 0.3 m and 0.5 m. Depths were obtained using a metal plate at the grass layer at the edge of the pit, whereby the lower side of the plate served as the zero reference level.

d. Incoming short wave radiation, vegetation canopy temperature, and soil moisture content

At the Cabauw site, 10-minute averages of the incoming short wave radiation fluxes are saved. Incoming short wave radiation is measured at the radiation field, which is located about 20 m away from the soil site. Incoming short wave radiation is measured at a height of 1.5 above the surface using a ventilated and heated Kipp & Zn CM11 pyranometer.

The vegetation canopy temperature is estimated from ten-minute averages of the infrared radiation temperature, as measured with a Heimann K-15-85 radiation thermometer. The thermometer is mounted at 2 m and it looks downward on the soil site with a M6 lens with a field of view of 40o. The field of view is thus a circle with a diameter of about 1.5 m, which is centered around the location where the soil temperature profile is taken.

Soil water content is measured at depths of 0.03 m and 0.08 m below the surface of the soil-terrain. It is measured using TDR-sensors manufactured by Campbell Scientific. In Cabauw, sensor type CS615 is used, which has a rod length of 0.3 m, while the width between the rods is 0.0032 m. For calibration the standard calibration function provided by the manufacturer is used. It should be noted that calibration for the (clayey) soil found at Cabauw is uncertain. Averages over 10 minutes are saved.

4. Setup of study

a. General numerical setup and assimilation window

We have run the (adjoint) model for 2003. As the soil type changes at a depth of 0.18 m below the surface, the bottom of the modeled soil column is chosen to be at 0.12 m ($z_b = 0.12$ m). Δz is set at 0.005 m. To optimize the involved parameters a data-assimilation

time window of 24 hours is used. Each 24 hour period starts at 0:00 UTC. The model is initialized by linear spatial interpolation of the first ten-minute averaged temperature profile after 0:00 UTC. Δt is set at 600 s, implying that $N = 144$. At the lower boundary, the model is forced by the measured ten-minute averages of the soil temperature at z_b. At the upper boundary of the soil column, the model is forced using the ten minute averages of the temperature measured by the Heimann thermometer, and ten-minute averages of the incoming short wave radiation flux.

b. Parameter prior values and uncertainty

Because a time-assimilation window of 24 hour is adopted, 24 hour averages for the control variables are required. As prior values for the control variables for each 24 hour period, their values on the previous 24 hour period are used. Prior uncertainties in these estimates are assumed to be uncorrelated. Off-diagonal elements of a priori error-covariance matrix are thus set to zero. Concerning the diagonal elements of the error-covariance matrix we assume that the error in each parameter is 50 % of the prior value: $C_p(1,1) = (\lambda_{pr}/2)^2$, $C_p(2,2) = (c_{p,pr}/2)^2$, $C_p(3,3) = (\Lambda_{pr}/2)^2$, and $C_p(4,4) = (\tau_{pr}/2)^2$.

At the beginning of the year 2003, the heat conductivity of the soil is initialized using the method as described by Johansen (1975) (see also Peters-Lidard et al. 1998). For the porosity of the B11 clay soil a value of 0.6 is adopted (Beljaars and Bosveld 1997), while the quartz content of the soil at Cabauw is assumed to be 0.1. The equation for the prior value of the soil heat conductivity at the beginning of 2003 then reads (Peters-Lidard et al. 1998):

$$\lambda_{pr} = 1.03\left(\log\left(\frac{\theta}{0.6}\right)+1\right)+0.13 \quad \text{W m}^{-1}\,\text{K}^{-1}, \tag{8}$$

where θ is the soil moisture content. Here, we estimate the soil moisture content as the average of the daily averaged soil moisture content at 0.03 m, and the daily averaged soil moisture content measured at 0.08 m. For the volumetric heat capacity, the prior value at the beginning of 2003 is calculated as:

$$c_{p,pr} = 0.4*2.0*10^6 + \theta*4.2*10^6 \quad \text{J m}^{-3}\,\text{K}^{-1}. \tag{9}$$

Note that the heat capacity of air has been neglected in this formulation. For Λ we adopt a prior value of 3 W m^{-2} K^{-1} at the beginning of the year, while τ is initialized at a prior value of 0.05 (Duynkerke 1992).

c. Applied measurements and uncertainty

Ten-minute averages of the soil temperature and soil heat flux measurements are available. As the first ten-minute average after 0:00 UTC is used for initialization, at each depth only 143 ten-minute averages of the soil temperature profile and the soil heat flux profile are independent measurements that can be used to optimize the parameters. Though the measurements can be assumed independent, modeling errors might be biased leading to time correlations between misfits between measurements and modeled equivalents (Bosveld and Bouten 2001). To mitigate this, we only use measurements taken at t=3:00 UTC, 6:00 UTC, 9:00 UTC, 12:00 UTC, 15:00 UTC, 18:00 UTC and 21:00 UTC, and 24:00 UTC. Thus, the observations obtained at each depth and at each time step are assumed independent and the off-diagonal elements of the matrix C_v are zero. The diagonal elements refer to the variance associated with the error in the soil temperature measurement and soil heat flux measurements that results from both instrumental errors and representation errors.

To calculate the variance of each ten-minute average of the heat flux through the plate, measurements at the three corners of the triangle are available. For each time step, however, only three observations are available. Therefore, for each ten-minute average, the variance is taken to be equal to the daily averaged variance, which is given by:

$$\begin{aligned} C_v(G05) &= \sum_{i=2}^{144} \frac{1}{286} \left(\left(GS05(i) - G05(i)\right)^2 + \left(GW05(i) - G05(i)\right)^2 + \left(GE05(i) - G05(i)\right)^2 \right), \\ C_v(G10) &= \sum_{i=2}^{144} \frac{1}{286} \left(\left(GS10(i) - G10(i)\right)^2 + \left(GW10(i) - G10(i)\right)^2 + \left(GE10(i) - G10(i)\right)^2 \right), \end{aligned} \tag{10}$$

where $C_v(G05)$ is the element of C_v that refers to the daily averaged variance of the plate heat flux measurements taken at approximately 0.5 m, $C_v(G10)$ is the element of C_v that refers to the daily averaged variance of the plate flux measurements taken at approximately 0.1 m, i denotes the index of the ten-minute averages; $GS05$, $GW05$, and $GE05$ denote the plate heat flux measurements taken at 0.5 m at respectively the Southern, Western and Eastern vertex of the equilateral triangle; $GS10$, $GW10$, and $GE05$ denote the plate heat flux measurements taken at 0.1 m at respectively the Southern, Western and Eastern vertex of the equilateral triangle; $G05$ denotes the average plate heat measurement at 0.5 m; $G10$ denotes the average plate heat flux measurement at 0.1 m.

For the soil temperature, only one profile is available. It is therefore not possible to estimate the variance of the soil temperature among different soil thermometers. To estimate the uncertainty in measured soil temperatures, data obtained by Scharringa (1976) are used. Scharringa measured the standard deviation of the measured soil temperature among 25 soil thermometers located at equidistant locations on a 18 m x 18 m grid, located in De Bilt in the Netherlands. He found that the standard deviation ranges from a value of about .3 K in winter to 1.1 in July. Here, we estimate the standard deviation of the soil temperature as:

$$C_v(T) = \left(.7 + .4 \sin\left(\frac{2\pi(dn - 104)}{365} \right) \right)^2, \tag{11}$$

where $C_v(T)$ are the elements of C_v that refer to the variance of the error in the soil temperature measurements at the different depths, dn is Day Of Year (DOY). Also, only temperatures taken at 03:00 UTC, 06:00 UTC, 09:00 UTC, 12:00 UTC, 15:00 UTC, 18:00 UTC, 21:00 UTC, and 24:00 UTC are used.

d. Minimization algorithm parameters

The steepest descent technique requires the specification of the step size matrix. It is calculated as (Margulis and Entekhabi 2004):

$$\gamma_q = \eta C_p, \tag{12}$$

where η is the scalar step size, which is set to a constant value of 1E-4. For the convergence factor, ε, we adopt a value of 1E-3.

5. Results and discussion

Our intent is to evaluate the performance of the proposed variational data-assimilation approach in providing estimates of G_0 as a term of the SEB. First, the approach is analyzed and evaluated by examining the characteristics of the cost function that it provides. Second, important side-products such as the optimal values for the model parameters are presented and discussed. Third, estimates of the surface soil heat flux obtained using the approach proposed in section 2 are compared to estimates of the surface soil heat flux obtained by the Lambda-approach. Last, the effect of using the proposed approach to estimate the surface soil heat flux on the SEB is explored.

a. Cost function characteristics

The approach as presented in section 2 has been applied for the entire year 2003. It appears that for 2003 there are 223 days for which all the data described in section 3 are available. For these days in Fig. 2 the posterior values of the cost function (J_{post}), i.e. the value of the cost function after optimizing is plotted against the prior value of the cost function (J_{prior}), i.e. the value of the cost function when it is evaluated using the prior values for the parameters. Also shown are values of 33.1 and 65.2, which define the 5 % and 95 % quantile of the Chi-square distribution with a number of degrees of freedom of 48.

From Fig. 2 it appears that on average J_{post} is about 70 % of J_{prior}. For values of J_{prior} lower than 35, the difference between the prior cost function and the posterior cost function is relatively small. In these cases, the observations contain only little extra information on the model parameters. For higher values of J_{prior} the difference between J_{prior} and J_{post} becomes much larger indicating that in these cases the observations put a much stronger constraint on the model parameters. Note that on some days, the reduction of J_{post} with respect to J_{prior} exceeds 50 %. As Margulis and Entekhabi (2004) argued, values of J_{post} on each day can be used to evaluate the validity of the prescribed error characteristics. To understand this, it is important to acknowledge that for independent observational errors, J_{post} is the sum of 48 (8x6+4-4) independent Gaussian distributions. As a result, it should theoretically be distributed according to a Chi-square distribution with 48 degrees of freedom (Menke 1989; Margulis and Entekhabi 2004). Such a distribution has a mean of 48 and a standard deviation of 9.8. If the posterior cost function is much smaller than 48, this suggests that either the prescribed observation error characteristics overestimate the uncertainty in the observations, or errors among different observations show correlations that are not accounted for. If the posterior cost function is much larger than 48, the prescribed uncertainty might be too small. Thus, Margulis and Entekhabi (2004) assessed the statistical confidence in the posterior parameter estimates and the statistical assumptions in the variational data-assimilation setup by evaluating whether the value of the posterior cost function was within one or two standard deviations away from the mean value. We follow a slightly different approach and validate the posterior cost function by evaluating whether it is within the 90 % confidence interval of its theoretical probability distribution. As such, posterior cost functions are only valid for

days in which the posterior cost function lies between 33.1 and 65.2. The 101 days for which this condition is met will be further analyzed, while the remaining days are not used.

For the selected days we show in Fig. 3 the contribution of different components to the entire posterior cost function (Eq. (2)) as given in Fig. 2. Components include a component penalizing deviations from the prior values, denoted by J_{ap} (Fig. 3a), a component penalizing deviations from the observed soil temperatures, denoted by J_T (Fig. 3b), and a component penalizing deviations from the observed soil heat fluxes, denoted by J_G (Fig. 3c). It appears that J_G dominates the posterior cost function throughout 2003. During the entire year, J_G scatters between a value of 20 and 60, with an average value of about 40. In contrast, J_T is smaller, ranging from a value of about 13 around DOY 2003 50 and 350 to a value of about 4 around DOY 2003 200. This is surprising as the data-assimilation cycle uses on each day 32 soil temperature measurements and 16 soil heat flux measurements. This suggests that the prescribed uncertainty in the soil heat flux as derived using Eq. (10) underestimates the real uncertainty in the soil heat flux measurements, while the prescribed uncertainty in the soil temperature measurements as derived using Eq. (11) overestimates the real uncertainty in the soil temperature measurements. Also, J_T has a clear seasonal cycle, with values of about 13 in winter conditions, around DOY 200 50 and 350, and values of about 3 in summer conditions, around DOY 2003 200. Interestingly, the observed seasonal cycle in the values of the posterior cost function, is reflected in the prescribed seasonal cycle in the standard deviation as inferred from Scharringa (1975). In winter conditions, Eq. (11) predicts relatively low values of the standard deviation leading to high values of the cost function or given discrepancies between observed soil temperatures and modeled equivalents. In summer conditions, Eq. (11) predicts relatively high values of the standard deviation leading to low values of the cost function or given discrepancies between observed soil temperatures and modeled equivalents. This suggests that the prescribed seasonal cycle as inferred from Scharringa (1975) overestimates the real seasonal cycle in the standard deviation.

J_{ap} is very small. Most points are very close to zero, though especially at the end of 2003 values of 0.8 are obtained. Apparently, the prior estimates only play a small role in determining the final estimates of the parameter values. Their prime role is to serve as a starting point for the iteration in algorithm 1.

b. Control parameter values

Fig. 4 shows posterior values of the soil heat conductivity, denoted by λ_{post} (Fig. 4a), the volumetric heat capacity, denoted by $c_{p,post}$ (Fig. 4b), the skin conductivity, denoted by Λ_{post} (Fig. 4c), and the transmissivity of the vegetation, denoted by τ_{post} (Fig. 4d). Also, in Figs. (4a) and (4b) values for λ and c_p are given that are calculated using respectively Eq. (8) or Eq. (9) and the average of the daily averaged soil moisture contents at 0.03 m and 0.08 m, The temporal evolution of the soil moisture at 0.08 m is not given as its calibration is uncertain. Its temporal evolution can be derived from the volumetric heat capacity as calculated using Eq. (9) as this quantity is linearly related to the soil moisture content.

It appears that from DOY 2003 50 to DOY 2003 230, λ_{post} scatters around 0.9 W m^{-1} K^{-1}. This value is comparable to the value derived from Eq. (8), though the latter shows somewhat less inter-diurnal variation. At about DOY 2003 230 λ_{post} decreases toward a value of about 0.6 W m^{-1} K^{-1}, around which value it scatters from DOY 2003 230 onwards. In contrast, the value for λ derived from Eq. (8) reaches a minimum value of about 0.8 W m^{-1} K^{-1} around DOY 2003 230, while it increases again to a value of about 1.0 W m^{-1} K^{-1} at DOY 2003 350.

The volumetric heat capacity follows a similar pattern as the soil heat conductivity. From DOY 2003 50 to DOY 2003 230, $c_{p,post}$ is relatively close to the value for c_p derived from Eq. (9). Both decrease from a value of about 3.3 MJ m^{-3} K^{-3} at DOY 2003 50 to a value of about 1.7 MJ m^{-3} K^{-3} at DOY 2003 230. From DOY 2003 210 onwards $c_{p,post}$ differs considerably from the value for c_p derived using Eq. (9). From DOY 2003 210 to DOY 2003 230, $c_{p,post}$ decreases further to a value of about 1.5 MJ m^{-3} K^{-3}, after which it increases sharply toward a value of about 2.7 MJ m^{-3} K^{-3} from DOY 2003 260 to DOY 2003 350. In contrast, the value for c_p derived from Eq. (9) reaches a minimum value of about 1.7 MJ m^{-3} K^{-3} around DOY 2003 230, after which it increases toward a value of about 2.5 MJ m^{-3} K^{-3} at DOY 2003 350.

Similar to λ_{post} and $c_{p,post}$, the temporal variation of Λ_{post}, shows a sharp change at DOY 2003 230. Before DOY 2003 230 Λ_{post} scatters around a value of about 4.0 W m^{-2} K^{-1},

whereas from DOY 2003 230 onwards, Λ_{post} scatters around a value of about 3.0 W m^{-2} K^{-1}. For τ_{post}, there is no clear regime change at DOY 2003 230. Throughout 2003 it decreases steadily from a value of about 0.03 at DOY 2003 50 to a value of about 0.005 at the end of 2003.

Parameter τ is only a function of the vegetation cover and the characteristics of the incoming short wave radiation. Parameters λ and c_p are mainly soil characteristics, while Λ is both influenced by the characteristics of the (top) soil, and the temperature distribution within the vegetation cover. Changes in the soil characteristics are thus the most probable explanation for the regime change that is observed for λ_{post}, $c_{p,post}$ and Λ_{post} at DOY 2003 230. A possible explanation for the regime change is the well-known property of a clay soil to crack in very dry conditions such as experienced during the summer of 2003 (Wallender et al. 2006). In very dry conditions cracking leads to the appearance of fissures in the clay soil that are filled with air which has both a very low conductivity and a very low volumetric heat capacity. Hence, the very low posterior values of the soil heat conductivity, the volumetric heat capacity and the skin conductivity from DOY 2003 230 to DOY 2003 250. From DOY 2003 250 onwards, the soil starts to wet again due to precipitation (note the increasing value of soil heat conductivity as derived from Eq. (8)). In these conditions, the fissures are filled with water which has a very large volumetric heat capacity, but a conductivity that is lower than the surrounding soil. Hence, the relatively large values of the volumetric heat capacity and the relatively low values of the soil heat conductivity and the skin conductivity from DOY 2003 250 onwards. It should however be noted that the current method assumes a homogeneous soil in which the vertical heat transfer can be described using the diffusion equation as given in Eq. (1a). Obviously, this assumption is violated when the clay soil cracks and another much more complex formulation should be applied to calculate the heat transfer within the soil column (Wallender et al. 2006).

c. Surface soil heat flux

In Fig. 5 ten-minute averages of G_0 estimated by the current approach are compared to ten-minute averages of G_0 estimated using the Lambda-approach as described by De Bruin and Holtslag (1982). Here, we have adapted the original approach described by De Bruin and Holtslag (1982) to account for the Philip (1961) correction (Eq. (7)) on the soil heat flux estimates. Unfortunately, unlike the new method, the Lambda method does not yield values

for λ that are representative for the soil column between the surface and 0.12. To account for the Philip correction, a constant "representative" value for λ of 0.9 W m^{-1} K^{-1} is adopted. Also shown in Fig. 5 is the 1:1 line. It appears that in conditions where G_0 derived using the new approach is between –40 W m^{-2} and –20 W m^{-2} (directed upward), the new approach gives lower values of G_0 than the adapted Lambda approach. When G_0 as estimated by the new approach is –40 W m^{-2} , while the adapted Lambda approach gives values of about –30 W m^{-2}. When G_0 is between –20 W m^{-2} and 30 W m^{-2} (directed downward), the new approach gives estimates of the surface soil flux that are similar to the estimates: in these situations the points in Fig. 5 scatter around the 1:1 line. When G_0 is larger than 30 W m^{-2} (directed downward), the new approach gives lower estimates of G_0 than the adapted Lambda method: when G_0 as estimated by the new approach is 50 W m^{-2}, the estimate obtained by the adapted Lambda approach is about 70 W m^{-2}.

It appears that the underprediction of G_0 by the new approach with respect to the Lambda approach can be mitigated when only daytime values of measured ground heat flux and soil temperature profile obtained are assimilated into the model. Similarly, the underprediction of the ground heat fluxes by the new approach with respect to the Lambda approach can be mitigated when only nocturnal values of measured ground heat flux and soil temperature profile are assimilated into the model. Furthermore, using only daytime measurements leads to somewhat larger values of the skin conductivity than using nocturnal profiles only. This suggests that especially Λ shows a diurnal cycle that favors large surface soil heat fluxes during day time, and decreases surface soil heat fluxes in nocturnal conditions.

The larger nighttime (negative) values of the surface soil heat flux as calculated using the novel approach result in a lower imbalance of the surface energy budget, as can be seen in Fig. 6 where we plot for both the adapted Lambda approach (a.) and the new approach (b.) G_0 against the measurements of (R_N-LE-H). It appears that for the Lambda approach, the estimated G_0 is not equal to (R_N-LE-H). When (R_N-LE-H) is negative, the difference amounts to a maximum value of 25 W m^{-2}, i.e. about 50 % of the value of (R_N-LE-H) at this point. For positive values of (R_N-LE-H) the difference between the estimated surface soil heat flux and the (R_N-LE-H) is about 100 W m^{-2}. The imbalance of the surface energy budget gets somewhat smaller when the novel approach to estimate the surface soil heat flux is used.

When (R_N-LE-H) is positive however, there is still a large imbalance between the measured (R_N-LE-H) and the determined surface soil heat flux.

6. Summary and conclusion

Our aim was to present and evaluate the performance of a new method for estimating the surface soil heat flux at the meteorological measurement facility at Cabauw in the Netherlands. The proposed method calculates the soil surface heat flux using an optimized land surface model that describes the transport of heat between the vegetation and the soil surface, the transport of heat through the interface between soil and air/vegetation continuum, and the transport of heat within the soil column. The model is optimized using a variational data-assimilation framework in which for each day optimal values for the parameters are calculated as a function of misfits between 48 uncertain measurements and simulated equivalents, and 4 uncertain prior parameter value estimates. Boundary conditions for the model are the (radiative) temperature of the vegetation, the temperature measured at 12 cm below the surface, and the incoming short wave radiation at the surface.

For the ground heat fluxes observational uncertainty is derived from the field measurements itself, while for the soil temperatures the observational uncertainty is assigned according to values for the lateral variations in the soil temperature found in literature (Scharringa 1975). Inserting the resulting values into the data-assimilation setup, it appears that the variational data-assimilation setup yields on 101 days values of the cost function that are between 33 and 67, the 90% confidence interval of a Chi square distribution with degrees of freedom is 48, the distribution that the cost function theoretically should follow assuming independent observational errors.

As important evaluation products, the proposed approach yields values of the model parameter: the soil conductivity, the volumetric heat capacity of the soil, the skin conductivity, and the transmissivity of the vegetation. Up to DOY 2003 230, values for the soil conductivity and the volumetric heat capacity follow theoretical values as derived from literature (Peters-Lidard et al. 1998). From DOY 2003 230 however deviations from the values derived from Peters-Lidard et al. (1998) start to occur, which is possibly due to the cracking of the clay soil in Cabauw.

It appears that in general the proposed method gives estimates of the soil heat flux that are in good agreement with the currently operational Lambda approach, provided that the

Lambda approach is adapted to include a term to account for a correction term that translates the heat flux over a heat flux plate to a soil heat flux. However, unlike the proposed approach which allows the incorporation of instrumental corrections that depend on the thermal parameters of the soil , the correction cannot be extracted by the method itself, but has to be quantified separately. Unfortunately, the proposed method gives very large estimates of the surface soil heat flux in conditions where the soil heat flux is larger than about 50 W m^{-2}, while it gives too small estimates when the soil heat flux is smaller than about -20 W m^{-2}. It appears that these discrepancies are caused by the lack of intra-diurnal variations in the skin conductivity in the proposed method. Therefore, it is recommended that the proposed approach is further refined so that the skin conductivity is allowed to vary within the day.

The conclusion is that the proposed approach is a good candidate to replace current operational algorithms to infer the surface soil heat flux from measured soil temperature and soil heat flux profiles. Its main advantages are that 1) the prescribed error statistics can be evaluated objectively by studying the distribution and characteristics of the cost function, 2) the model can incorporate different measurements that have different error characteristics, and 3) the model can include in a consistent way measurements that need corrections that depend on the modeled soil parameters themselves, as is the case for the soil heat flux plates. However, the current approach should be extended so that inhomogeneities of the soil as well intra-diurnal variations in the parameters can be accommodated.

Acknowledgements

This study is performed within the framework of the Climate Scenario 2 (CS2) project and the Mitigation Experiment 2 (ME2) project, which are both part of the BSIK Climate changes for Spatial Planning (CcSP) research programme. The authors would like to thank two anonymous reviewers for their useful comments.

Appendix

Eqs. (1a) to (1d) are solved using a finite difference technique. Finite differencing implies that the soil temperature is represented by its values at a discrete set of points in space and time. Discretizing the initial conditions yields:

$$T(k\Delta z, t_0) = f_1(k\Delta z), \quad k = 1,2,3,...M, \tag{A.1a}$$

where Δz is the vertical grid spacing, and $M = z_b / \Delta z$. For time integration of the values at discrete points in the vertical, the semi-implicit Crank-Nicolson technique is used, resulting in the following equation for the model given by Eqs. (1a) to (1d) (Press et al. 1986):

$$A\mathbf{T}(t_{i+1}) = B\,\mathbf{T}(t_i) + \frac{1}{2}\left(\mathbf{V}(t_{i+1}) + \mathbf{V}(t_i)\right),\ i = 0,1,2,...,N, \tag{A.1b}$$

where $t_i = t_0 + i\Delta t$, A and B are M x M matrices controlling the propagation of the soil temperature, Δt is the temporal grid spacing, N is the number of time steps, $\mathbf{T}(t_i)$ is a M x 1 vector with elements $\mathbf{T}(t_i)[k] = T(z_k, t_i)$, and $\mathbf{V}(t_i)$ is a M x 1 vector that contains information on boundary conditions. Matrices A and B are tridiagonal, implying that their elements are zero valued everywhere, except for the values on the main diagonal, the first diagonal below this and the first diagonal above this:

$$A(1,1) = 1 + \frac{1}{2}\left(\sigma + \Lambda\frac{\Delta t}{c_p \Delta z}\right),\quad B(1,1) = 1 - \frac{1}{2}\left(\sigma + \Lambda\frac{\Delta t}{c_p \Delta z}\right);$$

$$A(j,j) = 1 + \sigma,\ B(j,j) = 1 - \sigma,\quad j = 2,3,4,...M-1;$$

$$A(j,j-1) = A(j-1,j) = -B(j,j-1) = -B(j-1,j) = -\frac{1}{2}\sigma,\quad j = 2,3,4,...M-1;$$

$$A(M-1,M) = -B(M-1,M) = -\frac{1}{2}\sigma;$$

$$A(M,M-1) = A(M,M) = B(M,M-1), B(M,M) = 0,$$

where $\sigma = \lambda\Delta t / \left(c_p (\Delta z)^2\right)$. Elements of the vector $\mathbf{V}(t_i)$ are given by:

$$\mathbf{V}(t_i)[1] = \left(\Lambda T_v(t_i) + \tau K_{in}(t_i)\right)\frac{\Delta t}{c_p \Delta z};$$

$$\mathbf{V}(t_i)[j] = 0.,\quad j = 2, M-1;$$

$$\mathbf{V}(t_i)[M] = f_2(t_i).$$

Using Eqs. (A.1a) and (A.1b) as the forward model as a hard constraint, the discrete formulation of the Lagrange function becomes (Margulis and Entekhabi 2001,2004):

$$L_d = J + \sum_{i=1}^{N}\left(\boldsymbol{\mu}(t_i)\right)^T\left(A\mathbf{T}(t_i) - \mathrm{B}\,\mathbf{T}(t_{i-1}) - \frac{1}{2}\left(\mathbf{V}(t_i) + \mathbf{V}(t_{i-1})\right)\right), \tag{A.2}$$

where L_d is the discrete formulation of the Lagrange function, and $\boldsymbol{\mu}(t_i)$ is a M x 1 Lagrange multiplier vector.

The minimal value of the Lagrange function within the finite difference framework is obtained by setting the derivatives of the Lagrange function with respect to the state variables, the Lagrange multipliers and the model parameters to zero: $\left(\partial L_d / \partial \mathbf{T}(t_i)\right)^T = \mathbf{0}, \left(\partial L_d / \partial \boldsymbol{\mu}(t_i)\right)^T = \mathbf{0},\ i = 1,2,3,...,N$, and $\left(\partial L_d / \partial \mathbf{p}\right)^T = \mathbf{0}$.

Differentiating the Lagrange function with respect to the Lagrange multipliers, yields the discrete forward model as given by Eqs. (A.1a) and (A.1b). Differentiating the Lagrange function with respect to the soil temperature fields yields the discrete adjoint model, which is thus given by:

$$\left(\frac{\partial L}{\partial \mathbf{T}(t_i)}\right)^T = \mathbf{0} = A^T \boldsymbol{\mu}(t_N) + \left(\frac{\partial J}{\partial \mathbf{T}(t_N)}\right)^T; \tag{A.3a}$$

$$\left(\frac{\partial L}{\partial \mathbf{T}(t_i)}\right)^T = \mathbf{0} = A^T \boldsymbol{\mu}(t_{i-1}) - B^T \boldsymbol{\mu}(t_i) + \left(\frac{\partial J}{\partial \mathbf{T}(t_i)}\right)^T,\ i = 1,2,3,...,N-1. \tag{A.3b}$$

Finally, differentiating the discrete Lagrange function with respect to the model parameters yields:

$$\left(\frac{\partial L}{\partial \mathbf{p}}\right)^T = \mathbf{0} = \left(\frac{\partial J}{\partial \mathbf{p}}\right)^T + \sum_{i=1}^{N} \left(\left(\frac{\partial A\mathbf{T}(t_i)}{\partial \mathbf{p}}\right) - \left(\frac{\partial B\mathbf{T}(t_{i-1})}{\partial \mathbf{p}}\right) - \frac{1}{2}\left(\frac{\partial \mathbf{V}(t_i)}{\partial \mathbf{p}} + \frac{\partial \mathbf{V}(t_{i-1})}{\partial \mathbf{p}}\right) \right)^T \left(\boldsymbol{\mu}(t_i)\right) \tag{A.4}$$

References

Beljaars, A.C.M., and F.C. Bosveld, 1997: Cabauw data for the validation of Land Surface Parameterization Schemes. *J. Climate*, **10**, 1172-1193

Boni, G., D. Enthekhabi, and F. Castelli, 2001: Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *Wat. Res. Res.*, **37(6)**, 1713-1722

Bosveld, F.C., and W. Bouten, 2001: Evaluation of transpiration models with observations over a Douglas-fir forest. *Agric. For. Meteor.*, **188**, 247-264

De Bruin, H.A.R., and A.A.M. Holtslag, 1982: A simple parameterization of the surface fluxes of sensible and latent heat during daytime compared with the Penman-Monteith concept. *J. Appl. Meteorol.*, **21**, 1610-1621

Daley, R., 1991: Atmospheric Data Analysis, Cambridge Univ. Press, 457 pp.

Duynkerke, P.G., 1991: Radiation fog: A comparison of model simulations with detailed observations. *Mon. Wea. Rev.*, **119**, 324-341

Duynkerke, P.G., 1992: The Roughness Length for Heat and Other Vegetation Parameters for a Surface of Short Grass. *J. Appl. Meteor.*, **31(6)**, 579-586

Fuchs, M., and C.B. Tanner, 1968: Calibration and field test of soil heat flux plates. *Soil Sc. Soc. Am. Proc.*, **32**, 326-328

Garratt, J.R, and M. Segal, 1988: On the contribution of dew formation. *Bound.-Layer Meteor.*, **45**, 209-236

Heusinkveld, B.G., A.F.G. Jacobs, A.A.M. Holtslag, and S.M. Berkowicz, 2004: Surface energy balance closure in an Arid region: role of soil heat flux. *Agric. For. Meteor.*, **122**, 21-37

Jager, C.J., T.C. Nakken, and C.L. Palland, 1976: Bodemkundig Onderzoek van twee graslandpercelen Nabij Cabauw (in Dutch). NV Heidemaatschappij Beheer, 9 pp.

Johansen, O., 1975: Thermal conductivity of soils. Ph.D. thesis, University of Trondheim, 236 pp.

Lopez, P., A. Benedetti, P. Bauer, and M. Janiskova, 2006: Experimental 2D-VAR Assimilation of ARM cloud and precipitation observations. *Quart. J. Roy. Meteor. Soc.*, **617**, 1325-1347

Margulis, S.A., and D. Entekhabi, 2001: A coupled Land Surface-boundary layer model and its adjoint. *J. Hydrometeor.*, **2**, 274-296

Margulis, S.A., and D. Entekhabi, 2004: Boundary-layer entrainment estimation through assimilation of radiosonde and micrometeorological data into a mixed-layer model. *Bound.-Layer Meteor.* , **110**, 405-433

Massman, W. J., 1992: Correcting errors associated with soil heat flux measurements and estimating soil thermal properties from soil temperature and heat flux plate data. *Agric. For. Meteor.*, **59**, 249-266.

Mayocchi, C.L., and K.L. Bristow, 1995: Soil surface heat flux: some general questions and comments on measurements. *Agric. For. Meteor.*, **75**, 43-50

Menke, W., 1989: Geophysical Data Analysis: Discrete Inverse Theory. International Geophysics Series volume 45, 289 pp.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, 1986: Numerical recipes – the art scientific computing. Cambridge University Press, Cambridge, 818 pp.

Passerat de Silans, A., B.A. Monteny, and J. P. Lhomme, 1997: The correction of soil heat flux measurements to derive an accurate surface energy balance by the Bowen ratio method. *J. Hydrol.*, **188-189**, 453-465

Peters-Lidard, C. D., E. Blackburn, X. Liang, and E. F. Wood, 1998: The Effect of Soil Thermal Conductivity Parameterization on Surface Energy Fluxes and Temperatures. *J. Atmos. Sci.*, **55(7)**,1209-1224

Philip, J.R., 1961: Theory of heat flux meters. *J. Geophys Res.*, **66(2)**, 571-579

Portman, D.J., 1958: Conductivity and length relationships in heat-flow transducer performance. *Trans. Am. Geophys. Union*, **39**, 1089-1094

Scharrringa, M., 1976: On the representativeness of soil temperature measurements. *J. Agric. Meteor.*, **16**, 263-276

Sirkes, Z., and E. Tziperman, 1997: Finite difference of adjoint or adjoint of finite difference. *Mon. Wea. Rev.*, **125**, 3373-3378

Stull, R.B., 1988: An introduction to boundary layer meteorology. Kluwer Academic Publishers, Dordrecht, 666 pp.

Van de Wiel, B.J.H., R.J. Ronda, A.F. Moene, H.A.R. de Bruin, and A.A.M. Holtslag, 2002: Intermittent turbulence and oscillations in the stable boundary layer. Part I: a bulk model. *J. Atmos. Sc.*, **59**, 2567-2581

Van Ulden, A.P., and J. Wieringa, 1996: Atmospheric boundary layer research at Cabauw. *Bound.-Layer Meteor.,* **78**, 39-69

Van Wijk, W.R., and D.A. de Vries, 1963: Periodic temperature variations in a homogeneous soil. In: Physics in Plant environment, North Holland, Amsterdam, pp. 103-143

Wallender, W.W., K.K. Tanji, J.R. Gilley, R.W. Hill, J.M. Lord, C.V. Moore, R.R. Robinson, and E. C. Stegman, 2006: Water flow and salt transport in cracking clay soils of the Imperial Valley, California. *Irr. Drain. Sys.*, **20**, 361-387

Wösten, J.H.M., G.J. Veerman, and J. Stolte, 1994: Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland; de Staringreeks. (in Dutch), Technical Document 19, Winand Staring Centre for Integrated Land, Soil, and Water Research, Wageningen, 66 pp.

Figure caption list:

Fig. 1. Schematic representation of the system which model equations are given by eqs. (1a) to (1d). Also included are the positions of the observations

Fig. 2. Comparison of posterior values of the cost function versus prior values of the cost function for days on which all the available data are available (dots). Also shown are the 1:1 line (dashed line), and the 0.05 (dotted) and 0.95 quantile (solid) of a Chi square distribution which degrees of freedom is 48.

Fig. 3. Daily values of the components of the posterior cost function that penalizes deviations from the prior values (a.), the component of the cost function that penalizes deviations from the observed soil temperatures (b.), and the component of the cost function that penalizes deviations from the observed soil heat fluxes (c.) for the selected days.

Fig. 4. Posterior values of the soil heat conductivity, and values of the soil heat conductivity calculated according to Eq. (8) (a.), posterior values of the volumetric heat capacity of the soil, and values of the volumetric heat capacity of the soil calculated according to Eq. (9) (b.), posterior values of the skin conductivity (c.), and posterior values of the transmissivity of the vegetation (d.) for the selected days.

Fig. 5. Comparison of the surface soil heat as calculated using the new approach with estimates of the surface soil heat flux using the adapted Lambda approach. Also shown is the 1:1 line.

Fig. 6. The surface soil heat flux versus the net radiation minus the atmospheric transport (positive) upward for the new approach (a.), and the adapted Lambda approach (b.). Also shown is the 1:1 line.

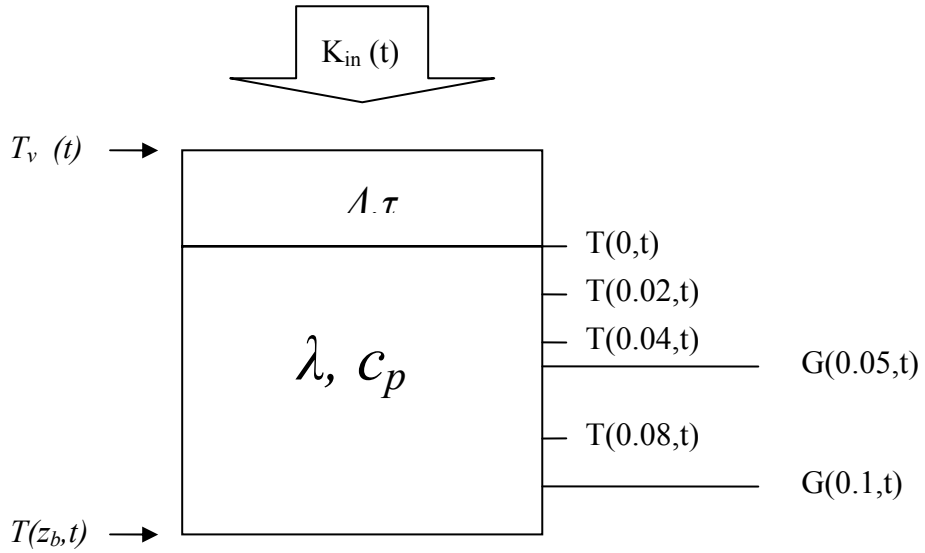


Fig. 1 Schematic representation of the system which model equations are given by eqs. (1a) to (1d). Also included are the positions of the observations

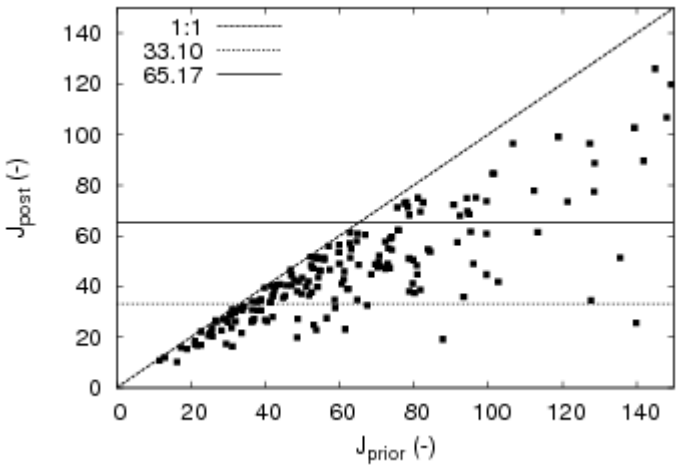


Fig. 2. Comparison of posterior values of the cost function versus prior values of the cost function for days on which all the available data are available (dots). Also shown are the 1:1 line (dashed line), and the 0.05 (dotted) and 0.95 quantile (solid) of a Chi square distribution which degrees of freedom is 48.

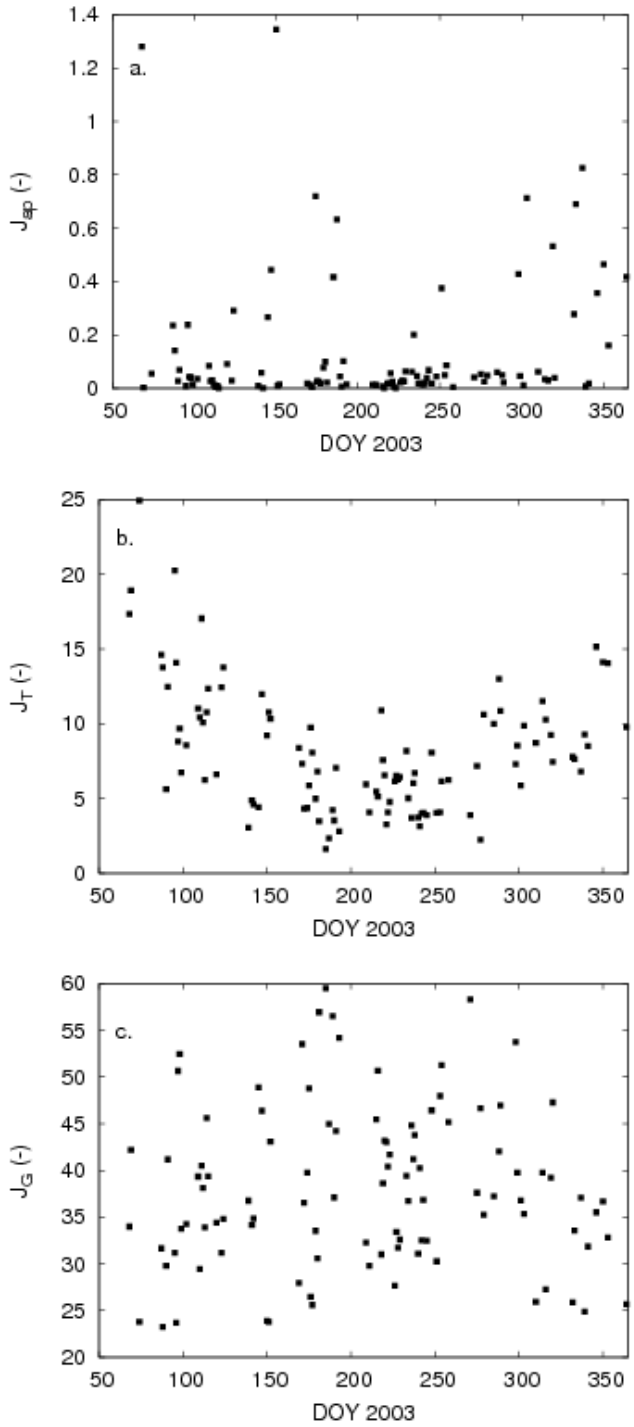
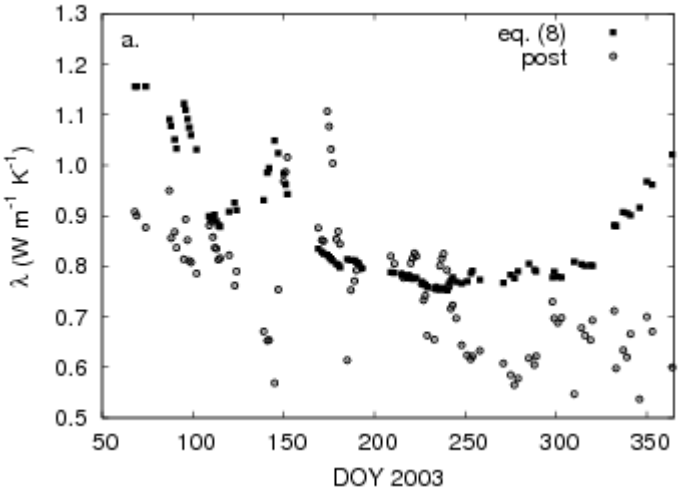
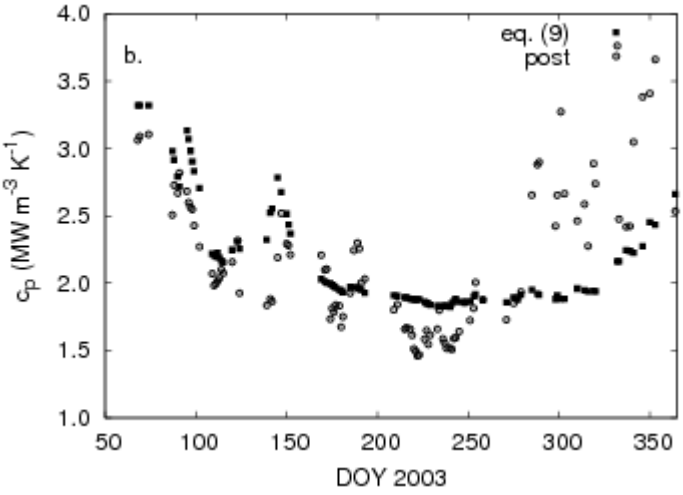
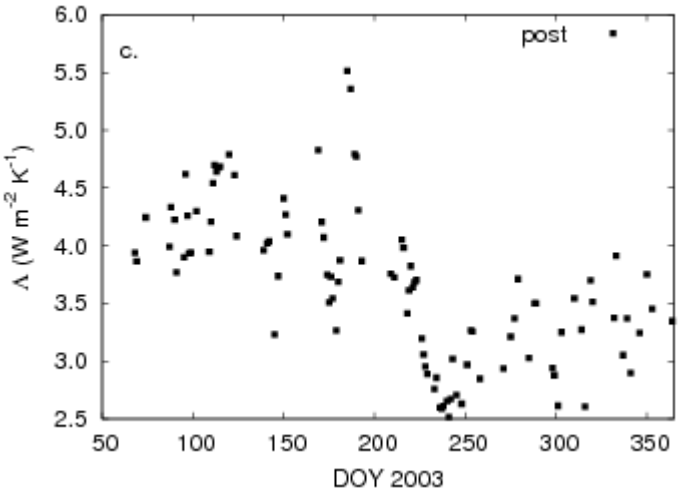


Fig. 3. Daily values of the components of the posterior cost function that penalizes deviations from the prior values (a.), the component of the cost function that penalizes deviations from the observed soil temperatures (b.), and the component of the cost function that penalizes deviations from the observed soil heat fluxes (c.) for the selected days.


a.
eq. (8)
post
λ (W m-1 K-1)
DOY 2003


b.
eq. (9)
post
cp (MW m-3 K-1)
DOY 2003


c.
post
Λ (W m-2 K-1)
DOY 2003

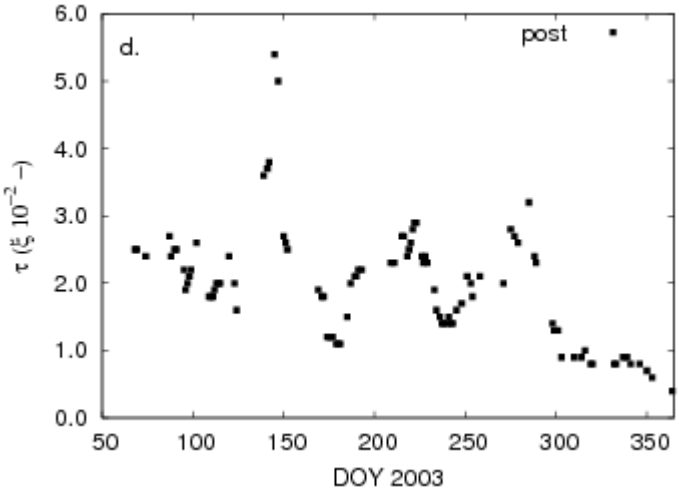


Fig. 4. Posterior values of the soil heat conductivity, and values of the soil heat conductivity calculated according to Eq. 8 (a.), posterior values of the volumetric heat capacity of the soil, and values of the volumetric heat capacity of the soil calculated according to Eq. (9) (b.), posterior values of the skin conductivity (c.), and posterior values of the transmissivity of the vegetation (d.) for the selected days.

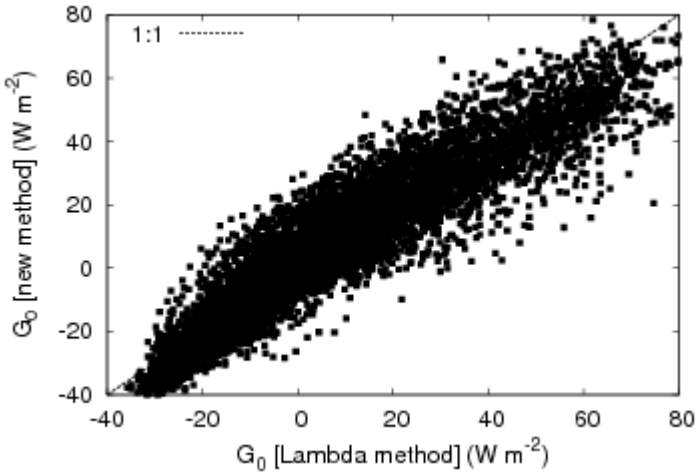


Fig. 5. Comparison of the surface soil heat as calculated using the new approach with estimates of the surface soil heat flux using the adapted Lambda approach. Also shown is the 1:1 line.

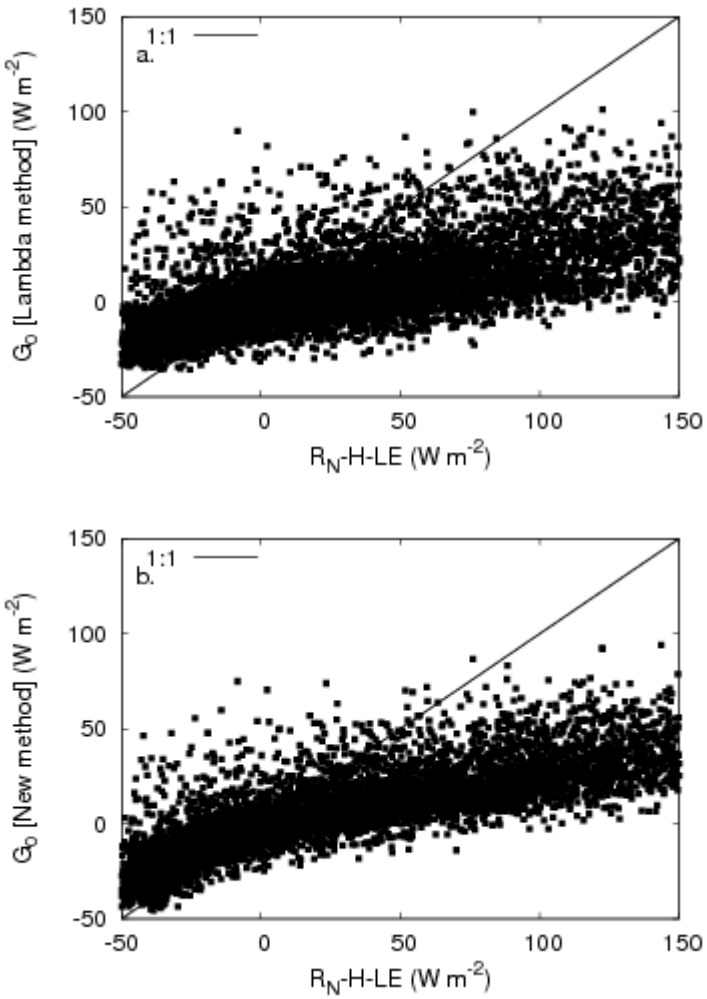


Fig. 6. The surface soil heat flux versus the net radiation minus the atmospheric transport (positive) upward for the new approach (a.), and the adapted Lambda approach (b.). Also shown is the 1:1 line.