

Representing land surface heterogeneity with the tiling method: merits and limitations in presence of large contrast

A. Manrique Suñén, A. Nordbo⁽¹⁾,
G. Balsamo, A. Beljaars,
and I. Mammarella⁽¹⁾

Research Department

⁽¹⁾ Department of Physics, University of Helsinki, Finland

July 2012

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



Series: ECMWF Technical Memoranda

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

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

Contact: library@ecmwf.int

© Copyright 2012

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

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

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

Abstract

The tiling method is used in many land surface models to represent the surface heterogeneity. Each gridbox is divided into fractions of different types of land use with independent solutions of the surface energy budget. An area-weighted average of the energy fluxes is computed to couple with the atmosphere assuming the air above the surface is well blended at a given height.

In the framework of validation of the tiling method, the ECMWF land surface scheme has been tested in offline mode driven by meteorological forcing provided by the ERA-Interim reanalysis. Two contrasting surfaces in the boreal region of southern Finland are considered: a Scots pine forest (Hyytiälä) and a small nearby lake (Valkea-Kotinen). The field observations are used to evaluate the land surface model simulations, for both energy fluxes and reservoirs. The model is able to characterize the main difference between the two sites, which appears in the energy partitioning, explained by the lake's large thermal inertia. In fact, while a large portion of the incoming solar radiation in the forest is released as sensible heat, the lake stores a substantial amount of energy in the water during late spring and summer and releases it in autumn. The different behaviour of the fluxes both on annual and diurnal timescales confirms the benefit of a tiling mechanism in presence of large contrast. However, it is also shown that the assumption of similar conditions at the blending height introduces errors.

1 Introduction

With the resolution of numerical weather prediction models constrained by computational cost, a single gridbox may span a wide variety of surface types. The sub-grid scale land surface heterogeneity must be parameterized in the surface scheme, so that the land characteristics are accounted for in the model. Two alternative approaches are normally used to represent different surface types within a gridbox. The first one combines the surface types to calculate effective parameters, and energy fluxes are derived from averaged gradients (e. g. Sellers et al. 1986). This method can be seen as a mixing of the different surfaces characteristics, where the surface flux resistances are arranged in parallel. In the second approach, or tiling method, suggested by Avissar and Pielke (1989) and Claussen (1991), each gridbox is divided into fractions of different surface types with independent solutions of the surface energy budget. The overall flux is an area-weighted average. Koster and Suarez (1992) and Essery et al. (2002) compared both strategies and showed similar results, the tiling method giving slightly lower turbulent fluxes, especially for contrasting surfaces. However, the former focused on heterogeneity due to different vegetation types, and in the latter the comparison is made against climatology. In the current study, we consider the more extreme situation of a forest and a lake, and concentrate on a local situation with observations for both surfaces.

The tiling approach has been widely adopted in land surface schemes for its simplicity and flexibility, because adding a new surface type to the system does not imply high computational cost. Limitations of the tiling concept are related to the imposition of a horizontally well mixed atmosphere above the different tiles at a certain height (blending height), which might not be valid for heterogeneities with large horizontal length scales (Koster and Suarez 1992).

The Hydrology Tiled ECMWF Scheme for Surface Exchanges (HTESSEL) is the land surface scheme used in the Integrated Forecast System, (Viterbo and Beljaars 1995, Van der Hurk et al. 2000, Balsamo et al. 2009). The surface characteristics of a gridbox are represented as sub-grid fractions of

each surface type. Over land, each gridbox is divided into six tiles: high vegetation, low vegetation, interception reservoir, snow on low vegetation and bare ground, high vegetation with snow beneath and bare ground. Water gridboxes are described by two tiles, the sea-ice fraction, and open water. Energy fluxes and skin temperatures are computed for each tile, and an area-weighted average is calculated for the gridbox to couple with the atmosphere. In the land surface scheme HTESSEL, a gridbox is either considered as land or as water depending if the fraction of water coverage is below or over 50%. Therefore, small sub-grid water bodies are disregarded. Recently, a ninth tile has been added to account for the lake fraction in inland terrain, and the lake model FLake (Mironov et al. 2008) has been incorporated to the surface scheme to reproduce lake processes (Dutra et al. 2010). The land surface scheme version which includes the lake model is called LAKEHTESSEL. The tiling method allows the representation of lakes even when they cover only small fractions of the gridbox. FLake model has been implemented and tested in NWP models like COSMO, HIRLAM, Météo-France, and UK Met Office (Mironov et al. 2010, Eerola et al. 2010, Salgado and Lemoigne 2010, Rooney and Jones 2010). Lakes have shown to have an important impact on the 2 m air temperature forecast in Finland by reducing up to 2 K the mean absolute error especially in autumn. This is due to a warming up of the near surface atmosphere because of the water thermal inertia (e.g. Balsamo et al. 2012). This warming effect of lakes was also reported in climate simulations over Europe in Samuelsson et al. (2010). Lake modelling results show consistency across different one-dimensional parameterizations for shallow lakes (Stepanenko et al. 2010).

In this study we evaluate the surface fluxes over a forest and over a lake and compare them against site observations. Data from Hyytiälä Scots Pine forest and Lake Valkea-Kotinen, southern Finland, are used. These datasets from two contrasting surfaces at close locations represent a good framework for evaluation of the tiling approach. The energy exchange dynamics of these ecosystems are rather different, due to large diversity in their surface characteristics (roughness length, skin conductivity, albedo; e.g. Beyrich et al. 2006). They represent a good example of the variety of characteristics that can be found in a single gridbox, and have to be computed together before being coupled to the same atmospheric column. The aim of this work is to evaluate the tiling system of HTESSEL by running the model in offline mode for two cases representing single tiles of high vegetation and of inland water. Both experiments are forced with the same meteorological conditions from ERA-Interim reanalysis (Dee et al. 2010). The blending height assumption is tested by forcing the runs with ERA-Interim data from different heights and with observed meteorological data.

In the following section we describe the observational sites and the model setup and forcing. In section 3 the results are presented, focusing on the energy fluxes and their seasonal and diurnal evolution. Also, the use of observed data as forcing and the model sensitivity to the blending height is analysed. Then, a comparison against observations of some specific surface variables of both forest and lake is presented in order to assess the quality of the model. Finally, conclusions are summarized in section 4.

2 Datasets and methods

2.1 Observation sites and measurements

Observational data were measured at two sites in Southern Finland, a Scots pine forest and a small boreal lake. The stations are about 80 km apart but the overall large scale meteorological conditions of both environments are very similar, making the observations suitable to test the tiling concept for contrasting surfaces. In this study several measurements have been used either for model comparison and/or as model forcing.

Hyytiälä forest station (61°51'N, 24°17'E, 179 m a.s.l.) is situated in a homogeneously vegetated Scots pine forest (*Pinus sylvestris*). The canopy height in the year of study was 16.5 m. Low vegetation is dominated by lingonberry (*Vaccinium vitis-idaea*), blueberry (*Vaccinium myrtillus*) and mosses. The forest floor is covered by a humus layer approximately 5 cm thick. The site is also known as SMEAR II (Station for Measuring forest Ecosystem - Atmosphere Relations), where continuous long-term observations of pools and fluxes in soil-ecosystem-atmosphere started already in 1995 (Vesala et al. 1998, Hari and Kulmala 2005). Specific humidity and air temperature were measured at 16.8 and 33.6 m height. Direct and diffuse solar radiation were measured by a pyranometer situated at a height of 18 m. Net radiation in the wavelength range 0.3–40 μm was measured at a height of 70 m. Latent heat flux, sensible heat flux and momentum flux were measured at 23.3 m using an eddy covariance system, consisting of a 3-D ultrasonic anemometer and an infrared gas analyser measuring turbulent fluctuations of wind velocity components, temperature and water vapour mixing ratio. The fluxes were calculated as 30 min block-averaged covariances between the scalars (or horizontal wind speed) and vertical wind velocity according to commonly accepted procedures (Aubinet et al. 2000, Mammarella et al. 2009). The ground heat flux was measured by heat flux plates situated 10 cm underground. The humus layer temperature was measured at 2 cm depth, and the soil temperature was measured at 2–5 cm depth of the mineral soil under the humus layer, both using a silicon temperature sensor. Humus layer moisture was measured at 2 cm depth and soil moisture was measured in the 5–25 cm layer. Additionally, snow depth was measured daily during the winter season. For more details on technical aspects see Launiainen (2010).

Lake Valkea-Kotinen (61°14'N, 25°03'E, 156 m a.s.l.) is a small lake of an area of 0.041 km² and an elongated shape. It has a mean depth of 2.5 m and the maximum depth reaches 6.5 m. The characteristics and energy balance of the lake are described in Nordbo et al. (2011). The lake is surrounded by tall forest which causes a strong channelling of flow along the lake (Vesala et al. 2006). The lake is a source of CO₂ with a mean annual flux of 77 g C m⁻² yr⁻¹ from eddy covariance measurements (Huotari et al. 2011). The measuring devices are on a floating raft, the height of measurements is 1.5 m above the water. Specific humidity was measured by an infrared light absorption analyser. Short wave radiation and temperature were measured by a Davis instruments weather station. Wind components were measured by a 3-D ultrasonic anemometer. A net radiometer was used to measure net radiation. Long wave downward radiation was calculated from net radiation, solar radiation and surface temperatures. Turbulent fluxes were obtained using the eddy covariance technique similarly to the forest site but using a 1 hour averaging time. Water temperature was

measured every hour at 13 depths (0.2, 0.5, 1.0, 1.5, 2.0, 2.25, 2.5, 2.75, 3.0, 3.25, 3.5 and 4.0 m) with a thermistor string on a buoy in the centre of the lake. Data from the thermistor string are available throughout the year whereas other measurements were conducted only during the open-water period. The total heat storage change in the lake water was calculated from the temperature time derivative at each depth, integrated for the whole water column, and multiplied by water density and heat capacity (Nordbo et al. 2011).

2.2 Offline Surface Model

Offline surface simulations were run using HTESSEL and LAKEHTESSEL for the Hyytiälä site for the period January to December 2006. In order to allow the model prognostic surface variables to stabilise, the simulation was repeated three times for the same year, using the previous iteration output as initial condition for the following run. This was sufficient even for the slow lake variables, such as bottom water temperature. Snow depth has been initialized at the beginning of the year with the observed value. The time step in the simulation is 30 minutes. The characteristics of the site (surface elevation, type of soil, terrain deviation) were taken from the operational T1279 description of the nearest model grid point to Hyytiälä forest observation site. The surface cover of the model gridbox is dominated by 75% of high vegetation, bare ground represents 8% and inland water 17%. However, to ensure that one type of surface is analysed for each run, the model was run for two configurations, one assuming only forest and another assuming only lake.

For the first case HTESSEL was used, and the tile representing the forest (high vegetation) is set to cover the whole gridbox, and low vegetation and lake fraction are fixed to zero. Evergreen needleleaf trees represent the dominant biome. For this type of canopy the model has a vegetation coverage coefficient set to 0.90, which means the residual is considered bare ground. In this experiment however, the vegetation coverage is set to 1 to avoid the bare ground tile. Therefore the only tile is high vegetation and when existing, high vegetation with snow underneath and interception reservoir. The snow and the interception reservoir tiles are dynamic and their tile fractions vary with snow depth and the precipitation intercepted by the canopy. The soil type is classified as medium texture.

In the second case, LAKEHTESSEL was run using the same forcing, but setting lake as 100% gridbox coverage. LAKEHTESSEL uses the lake model FLake, (Mironov et al. 2008) for the lake parameterization. FLake is a simple model based on a two-layer representation of the lake water temperature profile, and on the integral budget of energy for the two layers. The lake's top layer is well mixed and its depth is driven by convection or by wind. Below the mixed layer, the structure of the thermocline is based on the concept of self-similarity. The shape of the temperature-depth curve is determined by a shape function and bounded on the top and on the bottom by the mixed layer temperature and the bottom basin temperature. The shape function is approximated by a fourth degree polynomial of a dimensionless depth, and it depends on a shape factor. The model also parameterizes ice and snow cover, and the thermally active upper layer of bottom sediments using a similar approach for their temperature profiles. In LAKEHTESSEL, however, the bottom sediments are neglected and the snow is parameterized by HTESSEL. The prognostic variables are the mixed layer depth and temperature, mean water temperature, bottom water temperature, shape factor and ice cover thickness

and temperature. FLake's performance in LAKEHTESSEL was evaluated in Dutra et al. (2010). The lake-specific parameters are the depth of the lake and the extinction coefficient which were set to 4 m and 3 m⁻¹, respectively, for the present simulation. The ground heat flux as computed in the model accounts for ice processes and does not correspond to the lake heat change as calculated in the observations. In order to come to a fairer comparison, modelled lake water temperatures were interpolated for the same depths for which observations were available and the lake heat storage for the model was calculated in the same way it was calculated for the observations (Nordbo et al. 2011). Applying this method, the lake heat storage change only accounts for water storage, as it is calculated from the variation of lake water temperatures. This method assures that there is no heat storage during the ice-covered period, while lake temperatures are constant.

2.3 Forcing data

The model was forced with 3 hourly meteorological data from the lowest level of the ERA-Interim reanalysis (Dee et al. 2011). The horizontal resolution at which the forcing was extracted is T255 which corresponds approximately to a 80 km reduced Gaussian grid. In the vertical the reanalysis consists of 60 atmospheric levels. The forcing data used corresponds to the nearest neighbour grid point to Hyytiälä forest, which is located in (62.11°N, 23.91°E), 34 km away from the observational site and level 60 in the vertical which is approximately 10 m high. The description of the land surface for this gridbox is 77 % of high vegetation, 9 % of bare ground and 14 % of lake fraction.

The same forcing is used in both experiments, over lake and over forest. Since the forcing data come from a gridbox composed mainly of high vegetation, a larger error is expected in the lake run, because the data represents better the forest. Moreover, in the offline simulation there is no feedback of the surface into the atmosphere. To analyse the possible error introduced by the forcing, some experiments were run using observed data as forcing. Forcing data from both sites are available, although some values are missing. In order to be able to use observations of specific humidity, wind, temperature etc., the data were gapfilled. This was done with a linear regression between ERA-Interim data and observations. The missing values were filled by the corresponding ERA-Interim values corrected by the regression coefficients. To avoid erroneous values in winter, specific humidity was limited to the saturation specific humidity with respect to liquid or ice dependent on the forcing temperature. Long wave radiation was not available at the forest site, so ERA-Interim values were used. The height of the data measurements is 1.5 m for the lake site and for the forest site there are two sets of observations, at 16.8 and 33.6 m height.

The height at which the forcing is imposed is assumed to be the blending height or the level where the air is horizontally well mixed above the different tiles. The ERA-Interim's lower level is 10 m which in practice could be rather low. To test the impact of the forcing height on both sites, simulations were made for both cases with ERA-Interim forcing for the same grid point but at levels 59 (30 m) and 58 (60 m).

3 Results

The extensive set of observations available from both sites allows a good validation of the model output. We first analyse the energy fluxes, for which the comparison is twofold: (i) against observations to assess the model performance and (ii) the forest site versus the lake site to illustrate the difference in energy partitioning. Then, as further assessment of the physical validity of both schemes, some forest prognostic variables are compared against observations, and lake water temperatures and ice cover duration are compared against lake observations.

3.1 Energy fluxes

The surface energy balance in the forest can be expressed as:

$$R_n + H + LE = G \quad (1)$$

Where R_n is the net radiation, defined as the sum of downward and upward short and long wave radiation, H is the sensible heat flux, LE the latent heat flux and G the ground heat flux. For the lake, G needs to be substituted by ΔQ , the lake heat storage change. All fluxes are positive downwards and have units of W m^{-2} . The model's lake heat storage accounts for ice processes and the energy balance equation is closed; however, the calculated ΔQ to compare with observations does not include ice, and the balance is not closed in winter. The energy balance closure of the observed data for the forest and the lake is discussed in Launiainen (2010) and Nordbo et al. (2011) and on average tends to be below unity when computed as turbulent fluxes ($H+LE$) divided by available energy ($R_n - G$). This is commonly observed for eddy covariance measurements (Wilson et al. 2002).

To cover all time scales, the comparison is made for both seasonal cycle and diurnal cycle for 4 selected months.

3.1.1 Seasonal cycles

Modelled and measured energy fluxes are shown for both sites as 10 day averages in Fig. 1. In both lake and forest, the net radiation is very similar, as it is mainly driven by incoming solar radiation and both runs are forced by the same meteorological data. In winter time, solar incoming energy becomes very low and the main contribution to net radiation is the surface thermal radiation, resulting in net negative values. This is compensated in the forest by downward directed sensible heat flux. The sensible heat in the forest site follows a marked seasonal cycle, in winter the flux is directed downwards and the rest of the year the ground is heated by radiation and heat is transferred upwards, reaching up to -100 W m^{-2} in summer months. Typically, the sensible heat flux seasonal peak in Hyytiälä forest is reached in May-June (Launiainen 2010), but with 2006 being an exceptionally warm year high sensible heat flux was maintained until August. In the lake case, due to the water's high thermal inertia there is a delayed seasonal cycle of less amplitude, only reaching -25 W m^{-2} in July. Both HTESSSEL and LAKEHTESSSEL are able to represent very well these seasonal cycles, having correlation coefficients of 0.92 and 0.68, and root mean square errors of 20 W m^{-2} and 11 W m^{-2} ,

respectively. The observed high sensible heat flux values at the lake site at the beginning of May are a result of the lake water being very cold just after the ice break-up, while the air has already heated up.

The latent heat flux in the forest is low during winter, (about -5 W m^{-2} on average in January) and it increases gradually towards summer driven by the vegetation's evapotranspiration. Its measured maximum value is around -100 W m^{-2} in July. The model gives a lower peak value by about 20 W m^{-2} , but does an overall good representation of the seasonal cycle, with a correlation coefficient of 0.94. The measured latent heat flux over the lake is similar to the flux over the forest with a lower maximum occurring in July ($\sim 5 \text{ W m}^{-2}$ less). The model tends to overestimate the evaporation from the lake from June to October. This overestimation reaches up to 44 W m^{-2} during the summer period and result in a mean bias of 16 W m^{-2} for the open water period. This is the biggest error shown by the model, and it is partly explained by the too low specific humidity values in the ERA-Interim forcing, producing a high water vapour deficit, as well as by an excessive wind mixing in the proximity of the lake (discussed further in section 3.a.3).

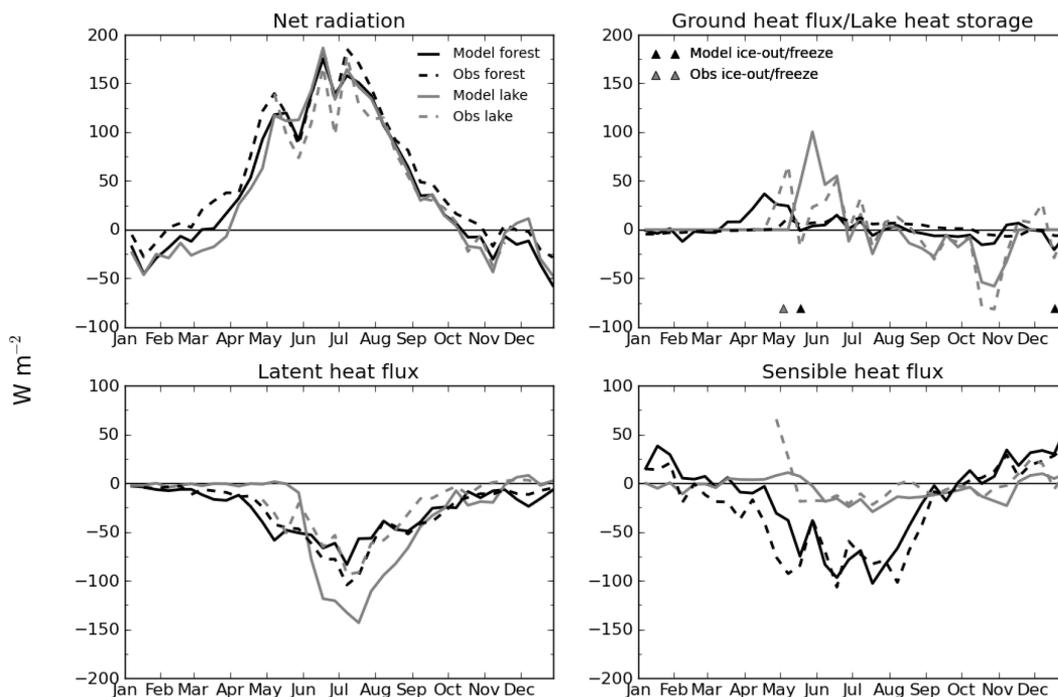


Figure 1. Seasonal cycle of 10 day averages of energy fluxes.

The measured ground heat flux at the forest site has a small negative value during winter time, as the soil is insulated by the snow layer. In spring as it melts, the ground starts to heat up. This is reproduced by the model, however, downwards heat flux is overestimated by about 25 W m^{-2} in April. For the lake, the energy storage change of the water mass for both observations and model is calculated from the water temperature time derivative as explained in section 2.a. Therefore, although ice processes are not accounted for, both model and observed values are comparable, and the heat storage change during the ice-covered period is zero. The beginning of the heat storage period in spring is retarded in the model, due to a delay of 14 days of the date of ice break-up. During this time, the lake accumulates energy that will be released during autumn.

In summary, the net radiation over both surfaces is very similar, and the main difference in the energy partitioning of a forest and a lake is that at a forest the energy is transformed roughly equally into latent heat and sensible heat, whereas at a lake the part which corresponded to sensible heat is stored in the water. Latent heat fluxes for both surface types are very similar in a seasonal scale as we are dealing with a case of non-stressed vegetation.

3.1.2 Diurnal cycles

The mean diurnal cycle based on hourly values of the fluxes have been plotted for January, April, July and October (Figs. 2-5). The net radiation diurnal evolution is very similar over the forest and over the lake, negative values at night and positive values at daytime mainly controlled by the incoming solar radiation. The systematic bias in winter is dominated by long wave cooling which is overestimated. This bias in ERA-Interim is seen in many models as documented by Wild et al. (2001). The peak is reached at 10 UTC (12 local time), and the amplitude of the cycle increases towards the summer months. From January to April the modelled net radiation during daytime at the lake site is lower than over the forest, which can be explained by the high albedo (up to 0.7 in the model) of the lake ice cover. However it cannot be verified with observations, as they are not available on the lake site until May.

The forest sensible heat flux follows a marked diurnal cycle. During the day there is an upwards heat flux and during night the flux reverses and has a lower magnitude. The diurnal cycle increases its amplitude towards the summer months (especially seen in July). It follows the net radiation cycle with opposite sign, and has its peak at 10 UTC (negative values - upward directed heat flux). HTESSEL represents very well the diurnal cycle. The amplitude of the diurnal cycle over the lake is rather small, as the surface water temperature does not vary as much as the land surface temperature, and the turbulent exchange coefficient for the lake tile is lower. During the ice cover period (December to April), turbulent fluxes are around zero and nearly all the radiative cooling is used to freeze the water. April's downwards heat flux is much lower than observed and most of this heat is used for the melting process. As a consequence the ice break-up is 14 days later in the model than observed. The reason for the underestimation of the sensible heat is not very clear. The surface temperature of 0°C must be correct, so we speculate that the Monin-Obukhov air/surface transfer formulation as used in the model is not very accurate for these transitions. This was also concluded by Nordbo et al. (2011). In summer, in contrast to the forest, the lake's maximum upwards heat flux occurs overnight (-27 W m^{-2} observed

at 2 UTC). Consequently, during night, while the lake heats the air above it, in the nearby forest the heat flux is going in the opposite direction, the ground being heated by the atmosphere. This may result in the generation of a local circulation at the interface of these contrasting surfaces, if their extension is big enough (Avisar and Pielke 1989). The effects of such a circulation are obviously not captured with the tiling concept. During August afternoons, small positive H values (heat flux towards the lake) were measured over the lake, corresponding to a stable stratification and the surface water temperature being about 1°C colder than the air (Nordbo et al. 2011).

Although the total latent heat fluxes throughout the year for forest and lake are quite similar (the forest evaporating slightly more overall), the diurnal patterns are rather different. In the forest, the evaporation comes mainly from the vegetation's evapotranspiration. It goes down to zero at night, and the maximum of -170 W m^{-2} occurs at 10 UTC in phase with the maximum solar radiation. The lake surface evaporates during the whole day in summer, from a minimum flux of -50 W m^{-2} at night to a maximum of -115 W m^{-2} two hours after the maximum radiation, resulting in a smoother and delayed diurnal cycle. The model performs very well for the forest. In the case of the lake, we can see that the excessive evaporation simulated by LAKEHTESSEL occurs especially during summer nights (discussed in the next section). The underestimation in April values can be explained by the delayed ice break-up in the model.

The ground heat flux is relatively small in the forest, and the model reproduces it well. The heat storage change in the lake has a marked diurnal cycle following the incoming solar radiation, the maximum storage calculated from observed data takes place at midday in July (283 W m^{-2}) corresponding to the maximum incoming solar radiation. The errors in heat storage change are of the order of 100 W m^{-2} in July and about 50 W m^{-2} during day time in October. This is non-negligible, but not unexpected because the errors in the modelled components of the surface energy balance accumulate in the heat storage. It is not straightforward to attribute errors to any component in the surface energy balance because the errors in the observed energy closure about 20-30% of the available energy for the monthly averages in the open water period in 2006 and 2007 (Nordbo et al. 2011).

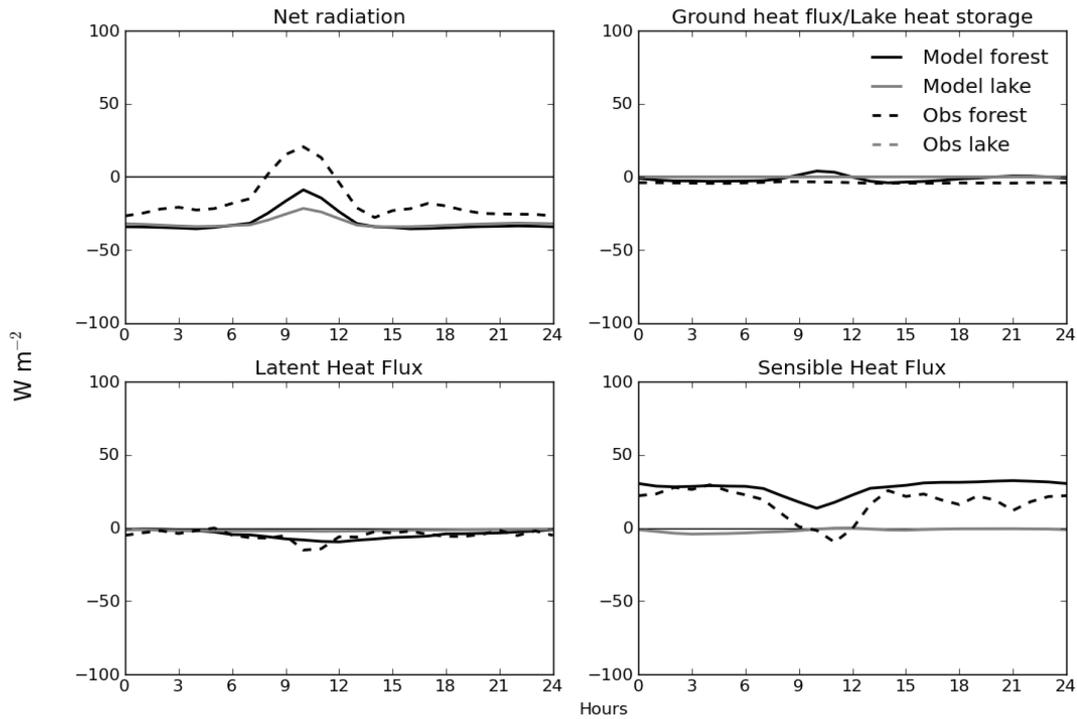


Figure 2. Mean diurnal cycle of energy fluxes for January. Note that measurements at the lake site are not available for this period.

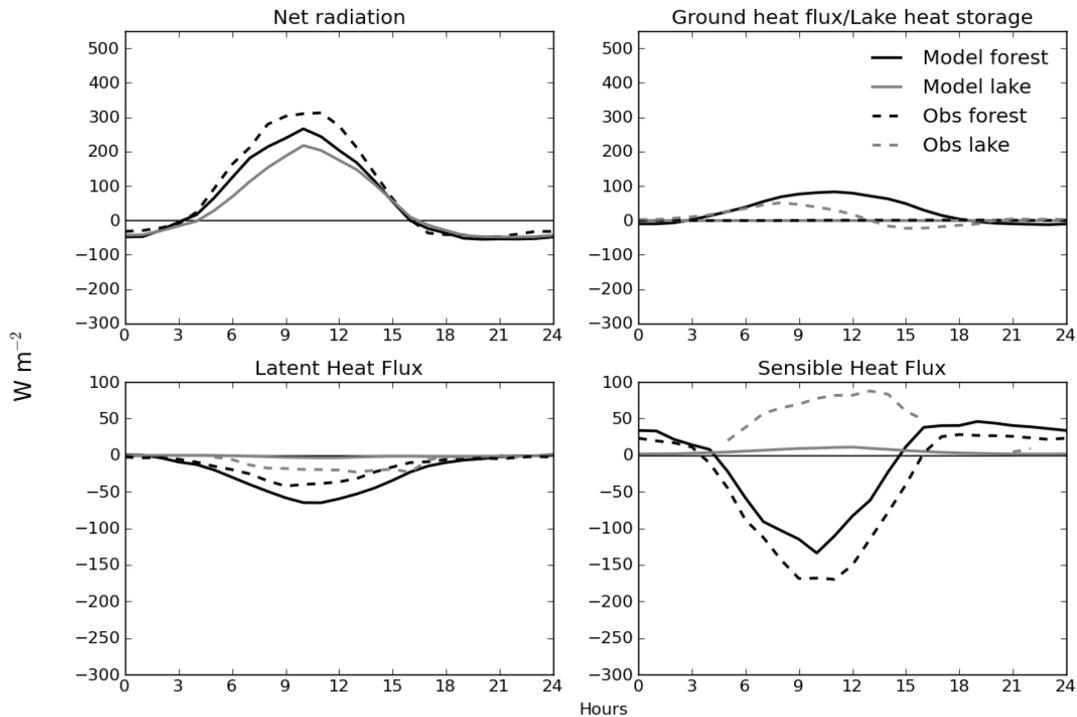


Figure 3. Mean diurnal cycle of energy fluxes for April. Note that net radiation measurements at the lake site are not available for this period.

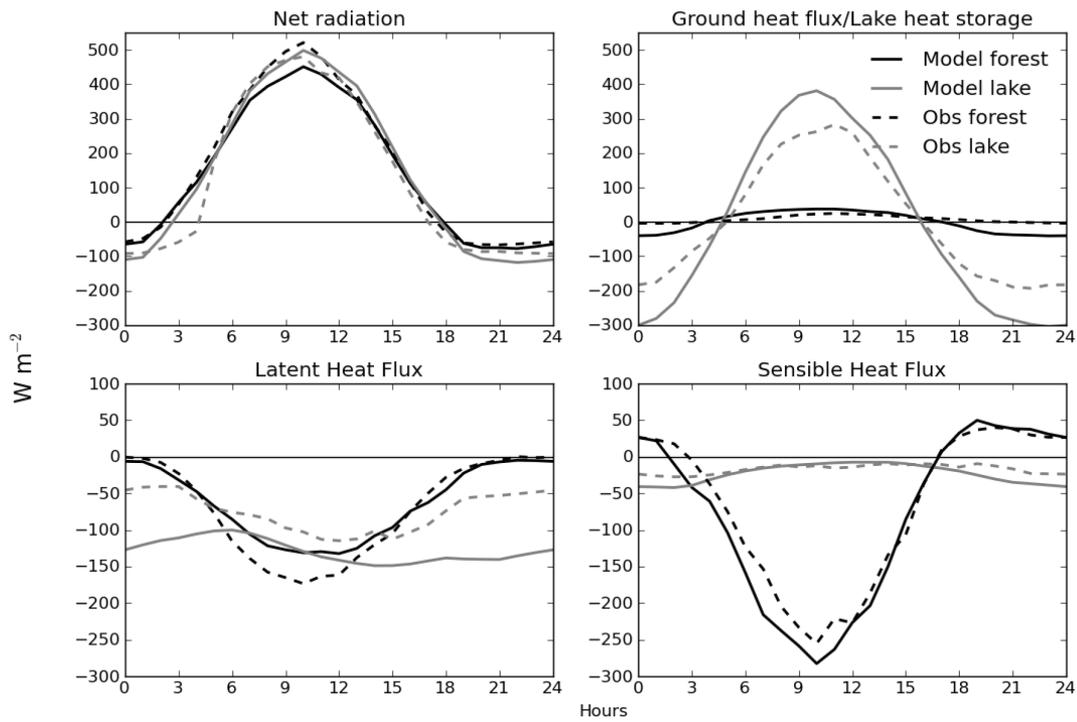


Figure 4. Mean diurnal cycle of energy fluxes for July.

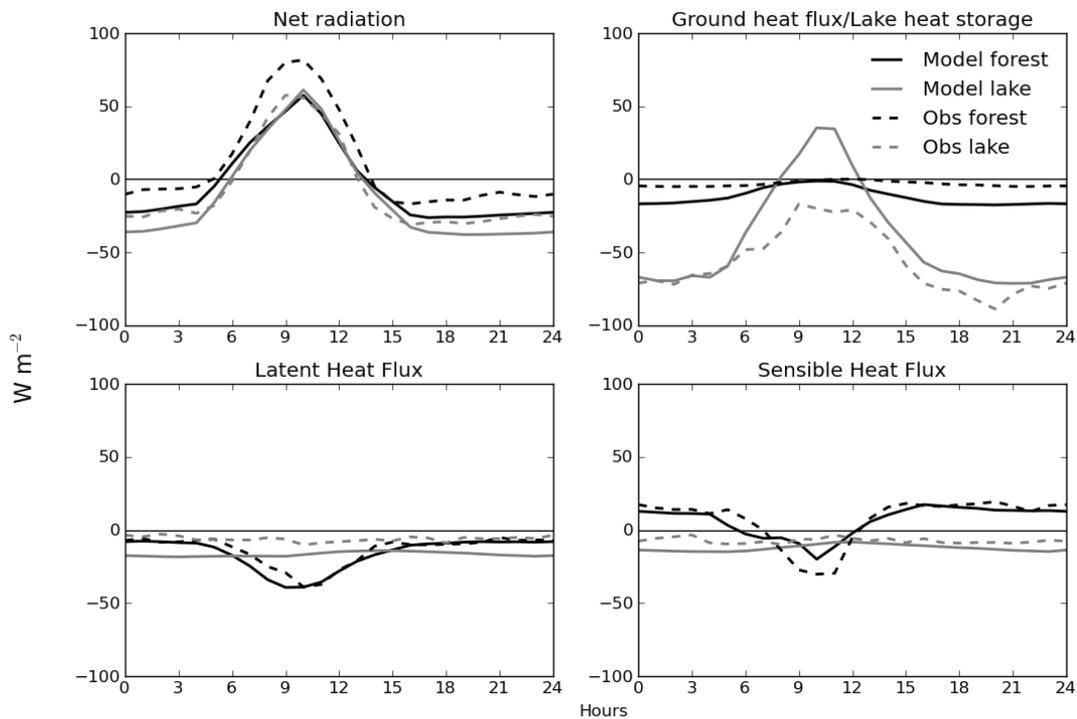


Figure 5. Mean diurnal cycle of energy fluxes for October.

3.1.3 Limitations of offline simulations and the use of observed forcing

In the offline simulations the same ERA-Interim forcing is used over both sites, which corresponds to the assumption of same conditions at the blending height as it would be presumed in a run over a tiled gridbox composed of both surfaces. Since the forcing data come from a gridbox composed mainly of high vegetation (section 2.c.), a larger error is expected in the lake run, because the data represents better the forest. The overestimation of latent heat over the lake could be a consequence of the ERA-Interim forcing not being fully representative of the lake conditions.

To analyse the error which might be introduced by ERA-Interim in the computation of latent heat flux by the model, some relevant parameters are studied. First, as the main driving factor for evaporation, the specific humidity deficit in both observations and ERA-Interim forcing is compared. Secondly, the horizontal wind speed is compared as it also plays a role in the latent heat flux, although Nordbo et al. (2011) found that LE over Lake Valkea-Kotinen was better explained by vapour pressure deficit alone than when multiplied by wind speed.

The specific humidity deficit for the observations is calculated as the difference between the saturation specific humidity corresponding to water temperature at 0.2 m depth (assumed as surface temperature) and the measured specific humidity at 1.5 m height. For the ERA-Interim data the specific humidity deficit is calculated as the difference between the saturation specific humidity for the modelled skin temperature and the forcing specific humidity at about 10 m high. The saturation specific humidity is calculated with respect to water or ice depending on the surface's temperature using Tetten's formula with the parameters set according to Buck (1981) for saturation over water and to Alduchov and Eskridge (1996) for saturation over ice. The annual cycle in Fig. 6 shows that the water vapour deficit introduced by ERA-Interim forcing when applied over a lake surface is too high compared to the real conditions during summer. This error can lead to the overestimation of latent heat flux by the model. However, in April and May the model's deficit is lower than observed, which together with the delay of ice break-up leads to an underestimation of latent heat flux. The minor negative values in winter indicate supersaturation. In the diurnal cycles (Fig. 7), in accordance with the seasonal plot, some negative values were found for the winter months. During summer the specific humidity deficit difference between model and observations is largest in the afternoons and at night, reaching values of around 2 g kg^{-1} in June, July and August (only July shown). This agrees with the overestimation of latent heat in summer which is more pronounced at night. It should be noted that in offline simulations the flux errors tend to be more pronounced than in coupled mode, because there is no feedback of the evaporation to the forcing humidity.

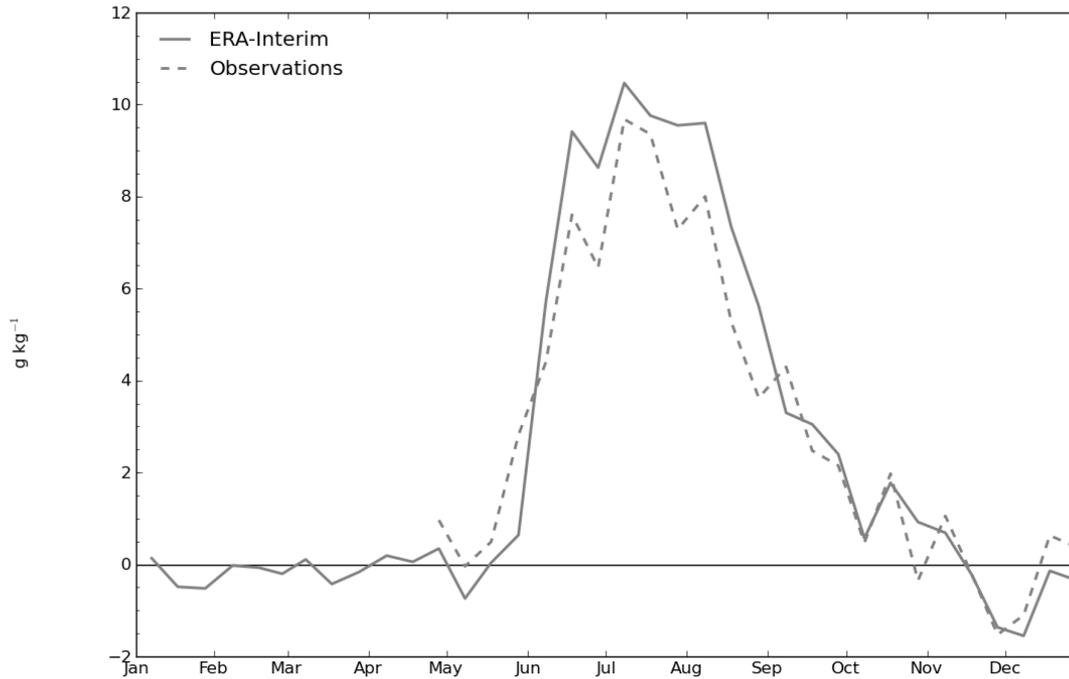


Figure 6. Seasonal cycle of 10 day averages of specific humidity deficit for the lake site.

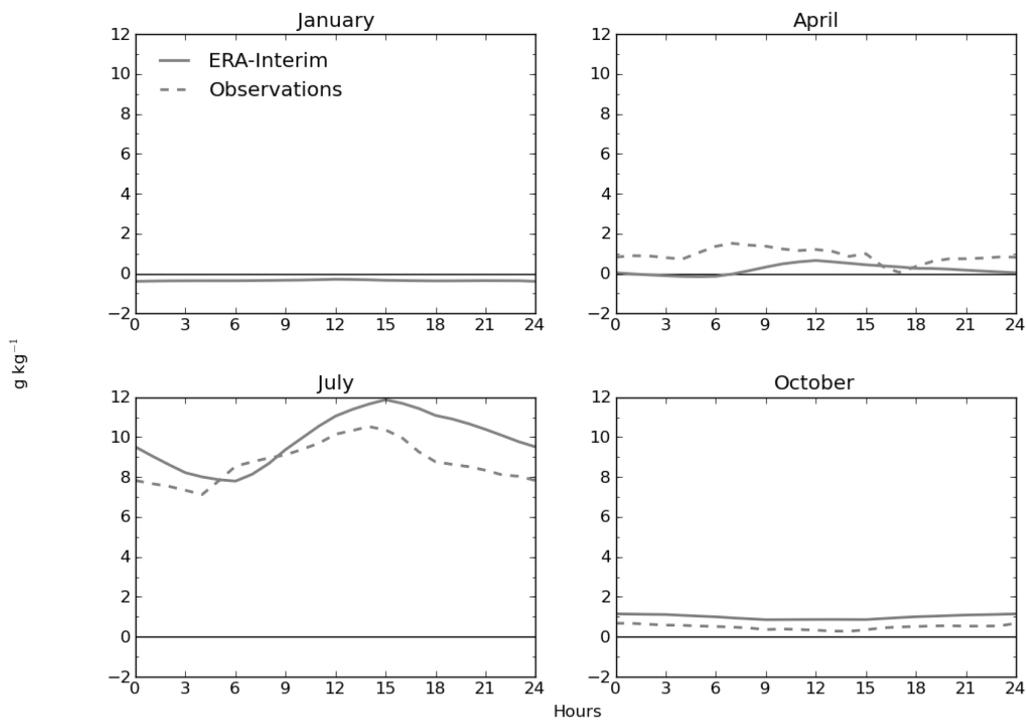


Figure 7. Mean diurnal cycle of specific humidity deficit for the lake site, for January, April, July and October.

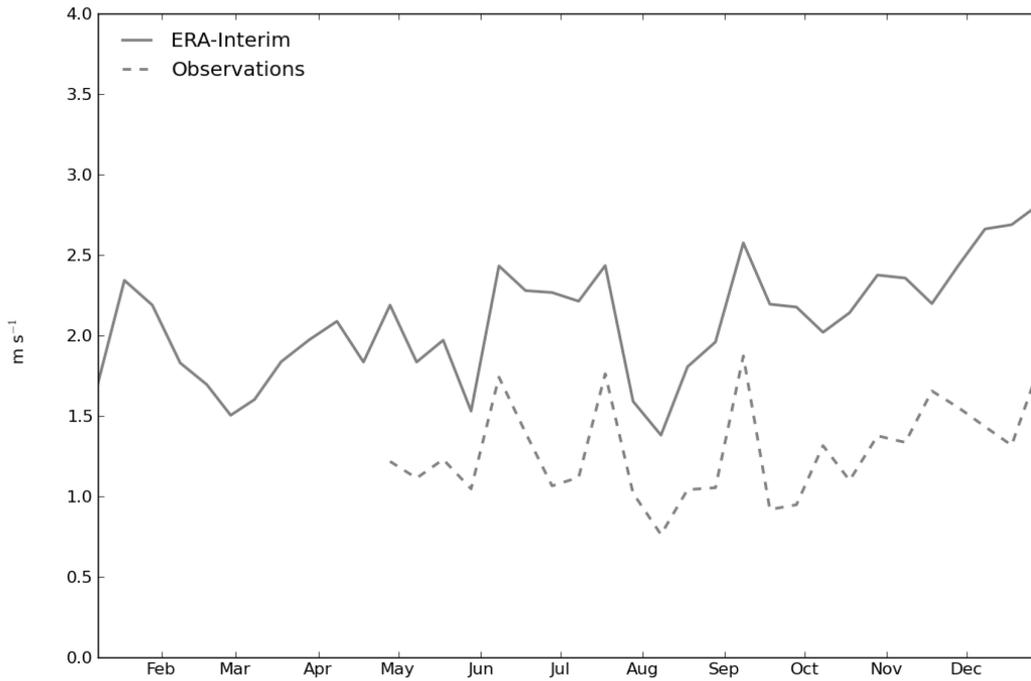


Figure 8. Seasonal cycle of 10 day averages of horizontal wind speed at 1.5 m for the lake site: interpolated ERA-Interim data and site observations.

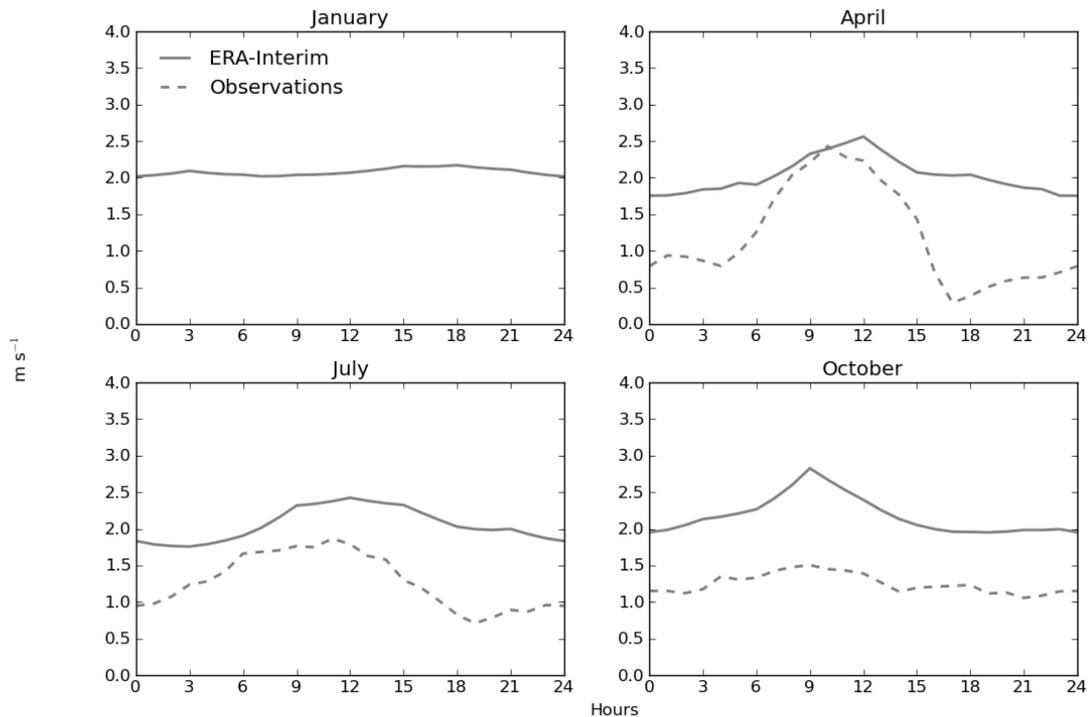


Figure 9. Mean diurnal cycle of horizontal wind speed at 1.5 m for the lake site for January, April, July and October: interpolated ERA-Interim data and site observations.

ERA-Interim reanalysis horizontal wind speed has been compared to observed values in Fig. 8 for the seasonal cycle and in Fig. 9 for the diurnal cycle. The ERA-Interim values at a height of 10 m have been interpolated to 1.5 m using Monin Obukhov similarity profiles to better compare against wind speed measurements of an anemometer on site at 1.5 m. ERA-Interim values are systematically higher than observed. Our lake of study is surrounded by forest, which act as a shelter, making the near surface air very calm. The observed and ERA-Interim wind diurnal cycles present a similar shape, with a maximum around midday and lower values at night, with observed winds always somewhat weaker.

LAKEHTESSEL was run using the observed data as forcing to evaluate how much of the excess in latent heat comes from the errors seen in the forcing horizontal wind and specific humidity. The latent heat flux obtained in this simulation is reduced, and the RMSE respect to the observed flux is decreased from 32 W m^{-2} when using ERA-Interim forcing to 19 W m^{-2} (Fig. 10). There is no longer the drastic increase of evaporation in June obtained with the ERA-Interim forcing, yet there is still an overestimation of $\sim 25 \text{ W m}^{-2}$ during the summer months. The reason may be that the model's transfer coefficients are not appropriate for this situation with highly inhomogeneous flow. The model's transfer coefficients over water are based on Monin-Obukhov similarity for homogeneous situations over the ocean. When analysing the diurnal cycles, we see that despite the latent heat values being reduced with the observed forcing, night evaporation is still overestimated, and the diurnal cycle is not well reproduced (July in Fig. 11). The highest errors appear at nighttime, which is when the lowest winds are registered, and stability effects (in terms of height divided by Monin-Obukhov length) on the transfer coefficients are greater (Beljaars 1997).

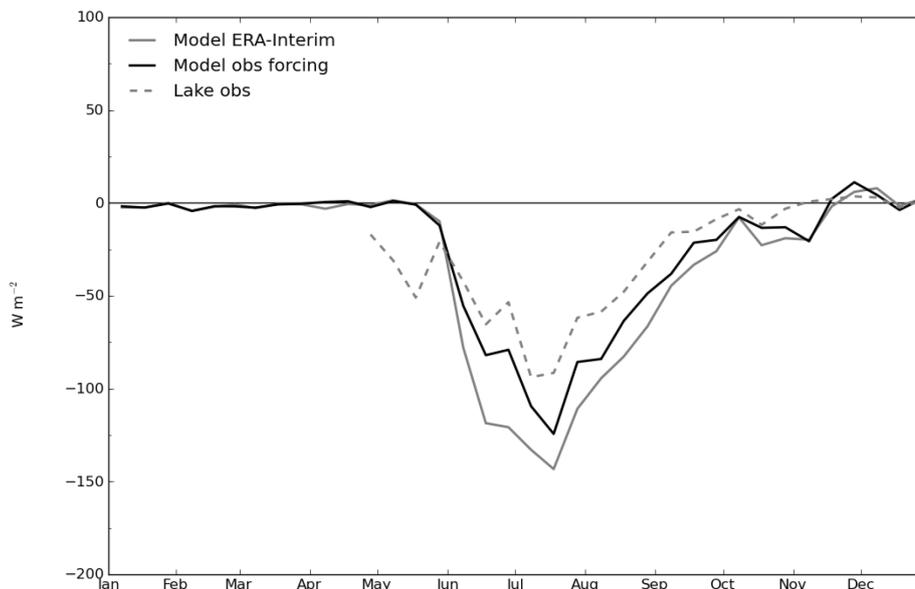


Figure 10. Seasonal cycle of 10 day averages of latent heat flux for the lake site: solid grey line for the model forced with ERA-Interim data, solid black line for the model forced with observed data, and dashed grey for the observations.

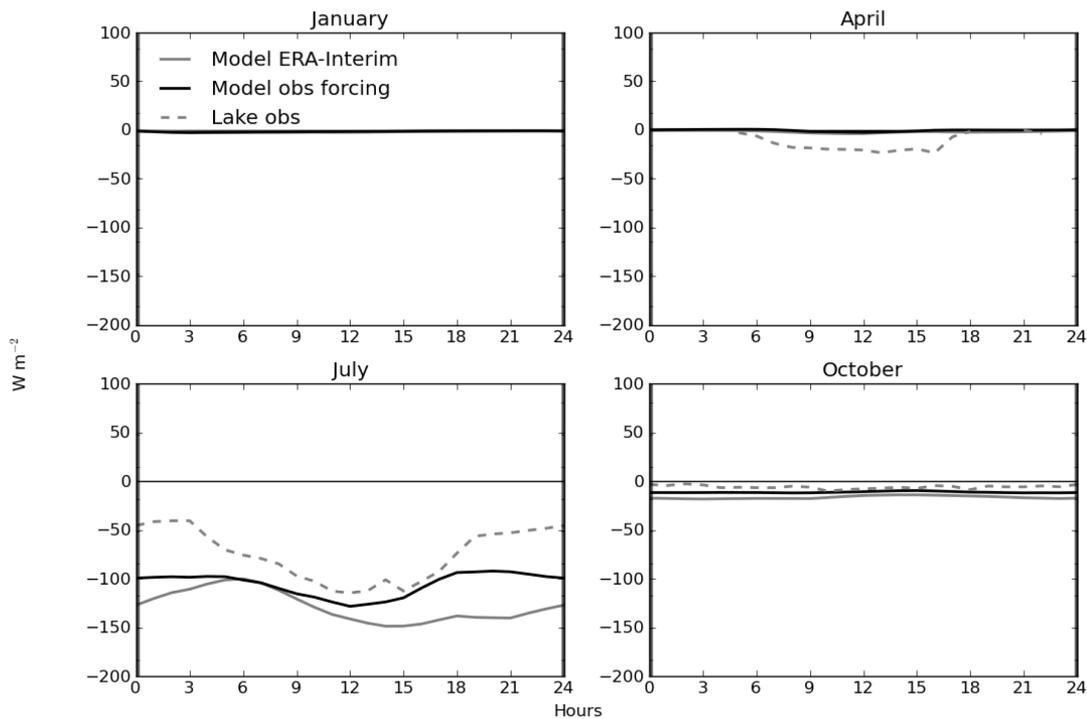


Figure 11. Mean diurnal cycle of latent heat flux for the lake sites for January, April, July and October: solid grey line for the model forced with ERA-Interim data, solid black line for the model forced with observed data, and dashed grey for the observations.

3.1.4 Forcing height variation

The lowest level in the ECMWF model (level 60 in ERA-Interim at about 10 m height) has been used as a forcing level for the individual tiles. Implicitly it is assumed to be the blending height where the internal boundary layers from the individual subgrid surfaces merge. In practice 10 m is rather low, so it might be advantageous to select a higher forcing level (or blending height).

To test the impact of the forcing height, simulations were made over the forest and lake sites using datasets from different heights as model forcing. From ERA-Interim reanalysis, higher levels were chosen: level 59 (30 m) and level 58 (60 m). To complete the study, observed measurements were also tested as forcing; these correspond to heights of 16.8 m and 33.6 m for the forest site and 1.5 m for the lake site. This allows not only a sensitivity test on both surfaces to the forcing height variation but also a comparison between ERA-Interim and observed forcing, as already partly examined in the previous section. The root mean squared errors of the energy fluxes obtained from these simulations compared to the observed fluxes are shown in Fig. 12.

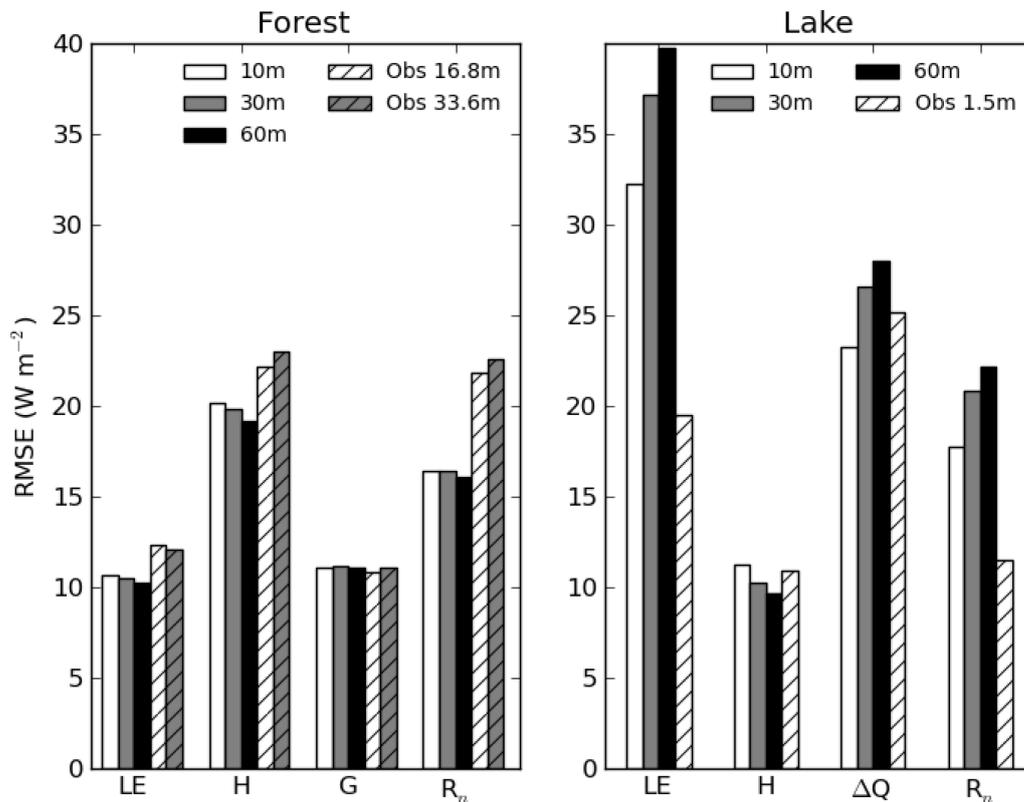


Figure 12. RMSE of energy fluxes obtained with ERA-Interim forcing at heights of 10, 30 and 60 m, and obtained with observed forcing at corresponding heights.

In the forest run, the effect of increasing the forcing height was slightly positive for all fluxes, reducing the RMSE by a few W m^{-2} , except for the ground heat flux where it was almost neutral. Overall, the effect of the forcing level over the forest is fairly small. The lake site however seems to be more sensitive to forcing height changes, although the results are mainly deteriorating. When increasing the forcing height, the evaporation from the lake tends to augment, worsening its RMSE by 7 W m^{-2} when using the 60 m height. It can be argued that, in the case of this lake environment being particularly calm as a consequence of the forest shelter, using forcing data from higher levels adds even more error in the winds, favouring evaporation. Sensible heat always tends to compensate the latent heat, giving a lower error. Higher forcing heights lead to more overestimation of the ice cover duration, and colder 2 m temperatures, due to less coupling between surface and atmosphere. It can be concluded that although an increased forcing height would be more compatible with the blending height concept, the improvements are not systematic.

When forcing with observed data the fluxes in the forest site present higher errors compared to the ERA-Interim forced runs (except ground heat flux, which does not vary much). The highest increase in error appears in the net radiation. This is surprising because since its most important component is incoming solar radiation, and this is imposed by the forcing, a better match with the observed values would be expected. However, the magnitude of the error is of the same order of an inconsistency found in the radiation measurements, probably due to uncertainties in the sensor calibration, (Pasi

Kolari, University of Helsinki, personal communication). For the year 2009, when also long wave radiation components have been measured at the site, the sum of measured incoming/outgoing short wave and incoming/outgoing long wave radiations is 9% lower than the net radiation as measured by a pyranometer at 70 m height. In the lake site a general improvement is obtained when using observed data as forcing, especially for the latent heat flux which reduces the RMSE almost by half (as already shown in section 3.a.3). This can be understood as the ERA-Interim data correspond to a grid box which is operationally composed of 77% of high vegetation, 9% bare ground and 14% of inland water. Therefore the data are more representative of the forest site, than of the lake site. Consequently, when using observed forcing, an improvement appears in the lake run, as the particular conditions of the site are better described by the observed data.

3.2 Forest soil temperature, moisture, and snow

The transfer of heat into the soil follows the Fourier law of diffusion (Hillel 1982). The model discretizes the soil into 4 layers, with the bottom of the layers at depths 0.07, 0.28, 1 and 2.89 m. The model soil temperatures represent the midpoint of each one of these layers. The ten day average values of the four model layers and observations for one layer are plotted in the top panel in Fig. 13. The upper layer has the biggest amplitude in the temperature cycle (20 K), as it is the one more linked to the surface. Its temperature varies from 273 K in February to 293 K in July. The second layer follows closely, being a few degrees warmer in winter and a few degrees cooler in summer. This cycle seems to be exaggerated if we compare to the milder observed soil temperatures measured between 0.02-0.05 m depths. The observed soil temperatures in winter time do not drop below 273 K due to the snow cover insulation, although the air temperatures reach 263 K. In summer the observed soil temperatures stay about 5-10 K lower than modelled. Despite the errors in soil temperatures it must be remarked that the 2 m air temperature produced by the model matches very well the observations, and has a seasonal amplitude of about 30 K, very similar to the model's first soil layer during the snow-free period. The deeper layers have a smaller and delayed cycle of temperatures, for example at 2.89 m depth the maximum temperature is reached in October.

The bottom panel of Fig. 13 shows 10 day average values of soil water content for the four model layers and measured values in the 0.05-0.25 m layer. The modelled soil moisture content follows an annual cycle with decreasing amplitude for deeper layers. The fourth layer (1-2.89 m) has an almost constant water content of $0.4 \text{ m}^3 \text{ m}^{-3}$. The measured values present larger amplitude in the cycle, underestimated by the model, with its maximum in May, right after the snow melting, and a minimum of about $0.1 \text{ m}^3 \text{ m}^{-3}$ in August. The humus layer appears to be constantly drier than the underlying soil, and follows a smaller cycle (not shown).

Snow depth was measured daily at the site. The model output is the snow mass which has been divided by the prognostic snow density to obtain a comparable snow depth (Fig. 14). The model snow has been initialized from the observed values on the first of January to better reproduce the snow layer (the other soil variables are iterated to equilibrium by repeating the annual cycle three times). The insulating effect of the snow in the ground is very important, without the initialization, there would be a upward ground heat flux causing a cooling of 10 K in the top soil layer in January and February.

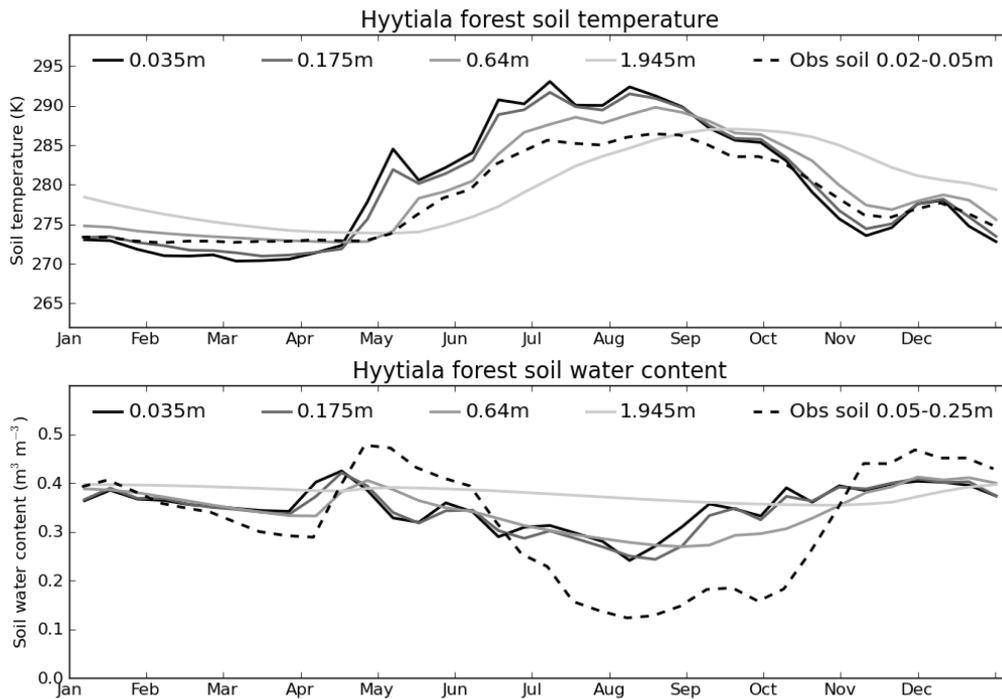


Figure 13. Seasonal cycle of 10 day averages of soil temperatures and water contents for the forest: solid lines for model output (depths refer to the middle of the layer), and dashed lines for observations.

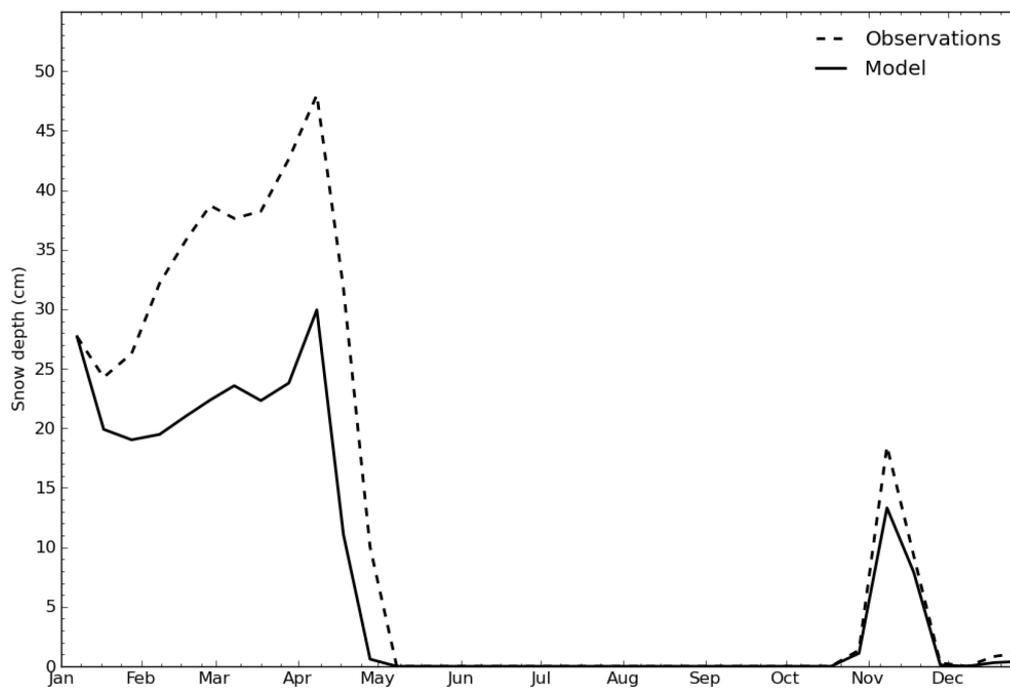


Figure 14. Seasonal cycle of 10 day averages of snow depth for the forest site.

3.3 Lake water temperature and ice cover

As most boreal lakes, Valkea-Kotinen is dimictic: the lake’s annual thermal cycle is marked by two mixing phases in autumn and spring, although the spring turnover is sometimes not complete (Huotari et al. 2009). In November or December, the surface freezes, and an inversion appears in the temperature profile, the upper part of the water being at freezing temperature and the bottom a few degrees warmer. After the ice melts, the water starts to be heated from the surface, forming a warm mixed layer which will deepen during summer. The turbidity of the water prevents the radiation from penetrating into deeper layers as has been discovered based on a Fourier analysis of water temperature (Nordbo et al. 2011). In October a turnover takes place and the column becomes isothermal.

Fig. 15 shows the water column evolution for the year of study, for the model and for the observed data. The colours indicate the water temperature and the grey line represents the ice cover thickness, there are no measurements of the ice depth, so it is set to 10 cm when there is ice. In the model, the temperature profile present at the moment of freezing is kept constant during the ice-covered period. All the column is at 273 K (the ice-water interface temperature), except for the deepest 0.5 m layer, which is 2 K warmer. This is slightly colder than observed temperatures, the surface temperature is correct, but the measured temperature profile presents a stronger inverse stratification. This difference appears because LAKEHTESSEL does not take into account the transfer of heat from the thermally

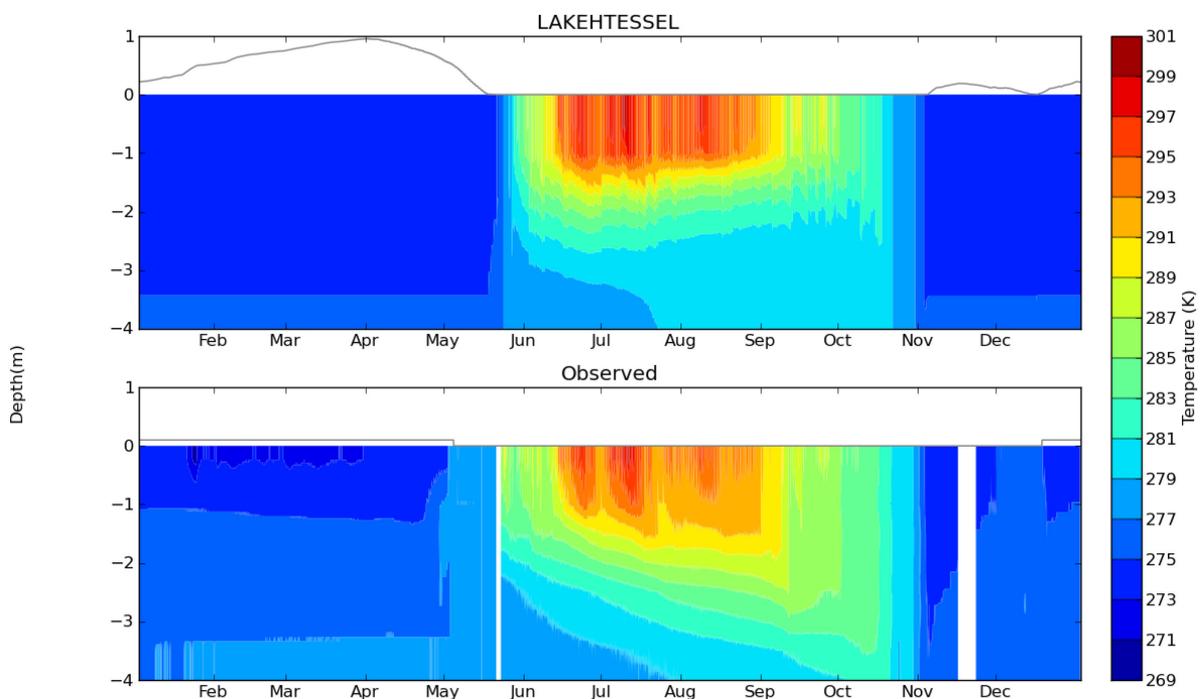


Figure 15. Seasonal cycle of lake water temperatures. The grey line indicates ice depth, there were no measurements of ice depth, therefore the grey line in the bottom panel only represents ice cover duration.

active upper layer of sediments. This heating is important in this lake because it is only 6 m deep, but for deeper lakes it would be negligible when compared to the remaining energy forcings acting on the lake (Dutra et al. 2010). After the ice break-up, the model fails to represent the deepening of the mixed layer depth, leading to an underestimation of temperatures at about 2.5 m depth. However it must be emphasized that the mixed layer water temperature calculated by the model matches very well the observations, with a root mean squared error of 1.23 K. This is the layer which is most directly linked to the atmosphere, so for numerical weather prediction purposes, the mixed layer water temperature is the most important prognostic variable from the lake model.

The observed ice break-up takes place on day 122 but in the model it is delayed by 14 days, while freezing occurs on day 355 and the model is 5 days early. This overestimation of the ice cover duration is most likely because LAKEHTESSEL does not model the snow on top of the lake ice. Consequently, the insulating effect of the snow layer is not accounted for and the too low temperatures produce a too thick ice layer and it is maintained for longer. This effect is analysed in Dutra et al. (2010). A threshold of 1 cm has been used to determine the existence of an ice cover in the model results.

4 Conclusions

The ECMWF Surface Scheme has been tested in offline mode for two very different surfaces forced by the same ERA-Interim reanalysis data. The availability of a good data set of observations from Hyytiälä forest and Lake Valkea-Kotinen sites has permitted the evaluation of the output fluxes and some specific variables. The model reproduces well the particularities of the energy partitioning of each site. It has been seen that both surfaces partition the incoming energy differently. The main difference is found in the partitioning into sensible heat and ground heat flux (or water heat storage change) in both seasonal and diurnal timescales. Due to the high thermal inertia of the water, the lake acts as an energy reservoir, and stores heat from the moment that the ice thaws in spring, until mid-summer, while in the forest the energy is released almost instantly as sensible heat. The latent heat flux is very similar in both cases. On a daily basis, for a typical summer day, similar differences are found, the radiative incoming energy over a lake is mainly stored in the water and over a forest transferred to the air in the form of sensible heat; with some energy used in latent heat dominated by the evapotranspiration cycle in the forest, and at a constant rate in the lake.

Although the difference between the tiling method and effective parameter method has been reported in literature to be fairly small, (Koster and Suarez 1992, Essery et al. 2002), it becomes more important when the surfaces are very contrasting. The different partitioning of energy for these two environments shows the need of separate subgrid surface energy balances to be able to describe the individual behaviours without mixing the available energy of tiles in a common energy balance equation. Moreover, the tiling method is conceptually clearer and more flexible.

However, the tiling method has also limitations since it is based on the assumption of the existence of a horizontal level, called the blending height, where the air is well mixed above the different surface components. The errors in evaporation over the lake are indeed partially the result of the forcing being more representative for a forest surface. In the case of contrasting surfaces and large heterogeneity length scales, the blending height is possibly higher than imposed by the model. The experiments with different forcing heights show that for the forest, errors in energy fluxes are slightly reduced when increasing the forcing height. The lake surface is more sensitive to the change in the forcing height, but the energy fluxes get worse when increasing the height. The reason for this worsening can be the especially calm conditions in the small lake.

The experiments run using observational data as forcing present an improved performance compared to ERA-Interim forcing runs only in the case of the lake. ERA-Interim data is more representative of the forest and the particular conditions for the lake site were better described by the observed data. The worsening of the results in the case of the forest might be due to imbalances in the observational dataset used as forcing.

Acknowledgements

We would like to acknowledge the support from the Academy of Finland through its Centre of Excellence program (project no 1118615), the Academy of Finland ICOS project (263149), the EU ICOS project (211574). We would also like to thank Dr. Jussi Huotari and Dr. Anne Ojala for conducting the measurements at Lake Valkea-Kotinen and Petri Keronen and Veijo Hiltunen, for maintaining the measurements at the Hyytiälä forest station and Dr. Samuli Launiainen for providing data from the forest site.

References

- Alduchov, O. A. and R. E. Eskridge, 1996: Improved Magnus form approximation of saturation vapor pressure. *J. Appl. Meteorol.*, **35**, 601-609.
- Aubinet, M., A. Grelle, A. Ibrom, Ü. Rannik, J. Moncrieff, T. Foken, A.S. Kowalski, P.H. Martin, P. Berbigier, Ch. Bernhofer, R. Clement, J. Elbers, A. Granier, T. Grünwald, K. Morgenstern, K. Pilegaard, C. Rebmann, W. Snijders, R. Valentini, T. Vesala, 2000: Estimates of the annual net carbon and water exchange of European forests: the EUROFLUX methodology. *Adv. Ecol. Res.* **30**, 113-175.
- Avissar, R., and R. A. Pielke, 1989: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.*, **117**, 2113-2136.

- Balsamo, G., P. Viterbo, A. Beljaars, B. van den Hurk, M. Hirschi, A.K. Betts, and K. Scipal, 2009: A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeor.*, **10**, 623-643.
- Balsamo, G., R. Salgado, E. Dutra, S. Boussetta, T. Stockdale, and M. Potes, 2012: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus-A*, **64**, 15829, DOI: 10.3402/tellusa.v64i0.15829.
- Beljaars, A.C.M., 1997: Air-sea interaction in the ECMWF model, ECMWF seminar proceedings on Atmosphere-surface interaction, 8-12 September 1997, 33-52, Reading.
- Beyrich, F., J.-P. Leps, M. Mauder, J. Bange, T. Foken, S. Huneke, H. Lohse, A. Lüdi, W.M.L. Meijninger, D. Mironov, U. Weisensee, and P. Zittel, 2006: Area-averaged surface fluxes over the LITFASS region based on eddy-covariance measurements. *Bound.-Layer Meteor.*, **121**, 33-65.
- Buck, A. L., 1981: New equations for computing vapor pressure and enhancement factor. *J. Appl. Meteor. Climatol.*, **20**, 1527-1532.
- Claussen, M., 1991: Estimation of areally-averaged surface fluxes. *Bound.-Layer Meteor.*, **54**, 387-410.
- Dee, D.P., S.M. Uppala, A.J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteor. Soc.*, **137**, 553–597, DOI:10.1002/qj.828.
- Dutra, E., V.M., Stepanenko, G. Balsamo, P. Viterbo, P.M.A. Miranda, D. Mironov, and C. Schär, 2010: An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal Env. Res.* **15**, 100-112.
- Eerola, K., L. Rontu, E. Kourzeneva, and E. Shcherbak, 2010: A study on effects of lake temperature and ice cover in HIRLAM. *Boreal Env. Res.* **15**, 130-142.
- Essery, R., M. Best, R. Betts, and P. Cox, 2002: Explicit representation of sub grid heterogeneity in a GCM land surface scheme. *J. Hydrometeor.*, **4**, 530-545.
- Hari, P. and M. Kulmala, 2005: Station for Measuring Ecosystem-Atmosphere Relations (SMEAR II). *Boreal Env. Res.*, **10**, 315-322.
- Huotari, J., A. Ojala, E. Peltomaa, J. Pumpanen, P. Hari, and T. Vesala, 2009: Temporal variations in surface water CO₂ concentration in a boreal humic lake based on high-frequency measurements. *Boreal Env. Res.*, **14**, 48-60.

- Huotari, J., A. Ojala, E. Peltomaa, A. Nordbo, S. Launiainen, J. Pumpanen, T. Rasilo, P. Hari, and T. Vesala, 2011: Long-term direct CO₂ flux measurements over a boreal lake: Five years of eddy covariance data. *Geophys. Res. Lett.*, **38**, L18401, DOI:10.1029/2011GL048753.
- Hillel, D., 1982: Introduction to soil physics, *Academic Press*.
- Koster, R., and M. Suarez, 1992: A comparative analysis of two land surface heterogeneity representations. *J. Climate*, **9**, 2551-2567.
- Launiainen, S., 2010: Seasonal and inter-annual variability of energy exchange above a boreal Scots pine forest. *Biogeosci.*, **7**, 3921-3940, DOI:10.5194/bg-7-3921-2010.
- Mammarella, I., S. Launiainen, T. Gronholm, P. Keronen, J. Pumpanen, Ü. Rannik, and T. Vesala, 2009: Relative humidity effect on the high frequency attenuation of water vapour flux measured by a closed-path eddy covariance system. *J. Atmos. Oceanic Technol.*, **26 (9)**, 1856-1866.
- Nordbo, A., S. Launiainen, I. Mammarella, M. Leppäranta, J. Huotari, A. Ojala, and T. Vesala, 2011: Long term energy flux measurements and energy balance over a small boreal lake using eddy covariance technique. *J. Geophys. Res.*, **116**, D02119, DOI:10.1029/2010JD014542.
- Mironov, D., 2008: Parameterization of Lakes in Numerical Weather Prediction Description of a Lake Model. *COSMO Technical report* No. 11.
- Mironov, D., E. Heise, E. Kourzeneva, B. Ritter, N. Schneider, and A. Terzhevik, 2010: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Env. Res.*, **15**, 218-230.
- Rooney, G. G., and I. D. Jones, 2010: Coupling the 1-D lake model FLake to the community land-surface model JULES. *Boreal Env. Res.*, **15**, 501-512.
- Salgado, R., and P. Le Moigne, 2010: Coupling of the FLake model to the Surfex externalized surface model. *Boreal Env. Res.* **15**, 231-244.
- Sellers, P.J., Y. Mintz, Y.C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation model. *J. Atmos. Sci.*, **43**, 505-531.
- Stepanenko, V.M., S. Goyette, A. Martynov, M. Perroud, X. Fang, and D. Mironov, 2010: First steps of a Lake Model Intercomparison Project: LakeMIP. *Boreal Env. Res.*, **15**, 191-202.
- Van den Hurk, B.J.J.M., P. Viterbo, A.C.M. Beljaars, and A.K. Betts, 2000: Offline validation of the ERA40 surface scheme. *ECMWF Tech. Memo.* **295**, 1-42.
- Vesala, T., J. Haataja, P. Aalto, N. Altimir, G. Buzorius, P. Keronen, T. Lahti, T. Markkanen, J.M. Mäkelä, E. Nikinmaa, S. Palmroth, L. Palva, T. Pohja, J. Pumpanen, Ü. Rannik, E. Siivola, H. Ylitalo, P. Hari, and M. Kulmala, 1998: Long-term field measurements of atmosphere-surface interactions in

boreal forest combining forest ecology, micrometeorology, aerosol physics and atmospheric chemistry. *Trends in Heat, Mass and Momentum Transfer*, **4**, 17-35.

Vesala, T., J. Huotari, Ü. Rannik, T. Suni, S. Smolander, A. Sogachev, S. Launiainen, and A. Ojala, 2006: Eddy covariance measurements of carbon exchange and latent and sensible heat fluxes over a boreal lake for a full open-water period. *J. Geophys. Res.*, **111**, D11101, DOI:10.1029/2005JD006365.

Viterbo, P., and A.C.M. Beljaars, 1995: An improved land surface parametrization scheme in the ECMWF model and its validation. *J. Climate*, **8**, 2716–2748.

Wild, M., A. Ohmura, H. Gilgen, J.-J. Morcrette, and A. Slingo, 2001: Evaluation of downward longwave radiation in general circulation models. *J. Climate*, **14**, 3227-3239.

Wilson, K., A. Goldstein, A. Falge, M. Aubinet, D. Baldocchi, P. Berbigier, C. Bernhofer, R. Ceulemans, H. Dolman, C. Field, A. Grelle, A. Ibrom, B. E. Law, A. Kowalski, T. Meyers, J. Moncrieff, R. Monson, W. Oechel, J. Tenhunen, R. Valentini and S. Verma, 2002: Energy balance closure at FLUXNET sites, *Agric. For. Meteor.*, **113**, 223-243.