BOREAL ENVIRONMENT RESEARCH 15: 100–112
ISSN 1239-6095 (print) ISSN 1797-2469 (online)

Helsinki 30 April 2010

An offline study of the impact of lakes on the performance of the ECMWF surface scheme

Emanuel Dutra[1)2)], Victor M. Stepanenko[3)], Gianpaolo Balsamo[4)], Pedro Viterbo[1)5)], Pedro M. A. Miranda[1)], Dmitrii Mironov[6)] and Christoph Schär[2)]

[1)] *CGUL, IDL, University of Lisbon, Faculdade de Ciencias Ed. C8, Campo Grande, PT-1749-016 Lisbon, Portugal*
[2)] *Institute for Atmospheric and Climate Science, ETH, Universitaetsstr. 16, CH-8092 Zurich, Switzerland*
[3)] *Moscow State University, Research Computing Center, Faculty of Geography, GSP-1, Leninskie Gory, building 4, RU-119991 Moscow, Russia*
[4)] *European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England*
[5)] *Instituto de Meteorologia, Rua C ao Aeroporto, PT-1749-077 Lisbon, Portugal*
[6)] *German Weather Service, Frankfurter Str. 135, D-63067 Offenbach am Main, Germany*

Received 15 Feb. 2009, accepted 5 May 2009 (Editor in charge of this article: Veli-Matti Kerminen)

Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P. M. A., Mironov, D. & Schär, C. 2010: An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal Env. Res.* 15: 100–112.

The lake model FLake was incorporated into the European Centre for Medium-Range Weather Forecasts (ECMWF) land-surface scheme HTESSEL. Results from global offline simulations are presented in order to evaluate the model performance in different climates and assess the impact of lake representation on the surface energy balance. The model was forced by ECMWF reanalysis ERA-Interim (1989–2005) near surface meteorology and surface radiation fluxes for the entire globe. Model validation includes lake surface temperatures and lake ice duration. The impact of the snow insulation on lake ice cover duration is discussed, as well as the sensitivity of results to the lake depth. Results show changes in the variation of surface energy storage, and changes in the partition of surface energy fluxes in high latitude and equatorial regions, respectively. General aspects concerning the incorporation of lake models into General Circulation Models for weather forecast and climate modelling are discussed.

Introduction

Lakes and other inland waterbodies can, in certain areas, constitute a large fraction of the land surface. As compared with vegetated land surfaces, inland waters play an important role in determining local and regional climates, primarily because of large differences in albedo, heat capacity, roughness, and energy exchange. Despite the radically different physical characteristics of inland waters as compared with their surrounding, most land surface models (LSMs) put more emphasis on the comparatively weaker differences within continental surface types (such as various types of vegetation and bare soil). Thus, so far sub-grid lakes have been largely neglected.

It has been shown that inland waterbodies can have important impacts on climate on various spatial scales (Bonan 1995, Hostetler *et al.* 1994, Krinner 2003, Lofgren 1997, Long *et al.* 2007). These studies report the role of open water surfaces in moistening the atmosphere, especially in summer (e.g. Bonan 1995). However, Lofgren (1997) reports that the Laurentian Great Lakes, by cooling and stabilizing the lower atmosphere in summer, tend to reduce evaporation and precipitation in the region. In winter, when boreal lakes are frozen, their climatic role is generally reported to be weak, although the energy flux through lake ice is significantly higher than over normal continental surfaces.

In this paper, the implementation of the lake model FLake (Mironov 2008) into the ECMWF LSM HTESSEL (Viterbo and Beljaars 1995, Van den Hurk *et al.* 2000, Balsamo *et al.* 2009) is presented. Offline (or stand-alone) simulations are often used to validate new parameterizations (e.g. Van den Hurk and Viterbo 2003) and in LSMs intercomparison initiatives (e.g. Dirmeyer *et al.* 1999, Rutter *et al.* 2009). The new reanalysis data from ECMWF ERA-Interim (Simmons *et al.* 2007), hereafter ERAI, is suitable for such a modelling framework. The reanalysis is available since 1989 to present, and in this study is used for 17 years (1989 to 2005). Global stand-alone simulations allow the verification of the lake model and behaviour of the proposed implementation in contrasting climate regions. Long-term simulations (17 years) also allow inter-annual verification and impacts analysis, since the model is exposed to different weather regimes.

Validation of global simulations requires independent global datasets. The lake surface-water temperature (LSWT) derived from the MODIS instrument on-board the TERRA platform is suitable for the validation exercise. Lake ice cover duration is an important feature of many high latitude lakes, but remotely sensed LSWT only allow a clear validation in ice free conditions. The Global Lake and River Ice Phenology (Benson and Magnunson 2007) provides a comprehensive dataset of formation and breakup of lake ice on several hundreds of lakes in the northern hemisphere.

The snow insulation effect has been reported as an important mechanism controlling energy exchanges between soil and atmosphere (Cook *et al.* 2008). This energy exchange governs the lake ice growth (or reduction) and subsequently affects the lake surface temperature. Simulations with and without snow over lake ice are analysed and the snow insulation effect is discussed. The lake depth is a crucially important model parameter, but there is not much relevant information regarding this parameter at the global scale. The sensitivity of the present validation to lake depth is also discussed. From the point of view of the atmosphere, the representation of lakes will change the surface fluxes. The impact of the representation of sub-grid scale lakes is evaluated on monthly time scales for different climatic regions.

Methods

Models

The LSM HTESSEL (Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land) represents vertical transfers of water and energy using four vertical layers to represent soil temperature and moisture (Viterbo and Beljaars 1995). The model evaluates the soil response to the atmospheric forcing, and estimates the surface water and energy fluxes and the temporal evolution of soil temperature and moisture content. At the interface between the surface and the atmosphere, each grid-box is divided into fractions (tiles), with up to six fractions over land (bare ground, low and high vegetation, intercepted water, shaded and exposed snow) and two extra tiles for open water (ocean) and sea-ice. Each fraction has its own properties defining separate heat and water fluxes used in the energy balance equation solved for the tile skin temperature. The model is part of the integrated forecast system (IFS) at ECMWF in applications ranging from short range weather forecast to reanalysis such as ERA-40 (Uppala *et al.* 2005) and ERAI. Along with its operational use, the model has served as a research test bed for a variety of applications ranging from leaf area index assimilation (Jarlan *et al.* 2008) to drought characterization (Dutra *et al.* 2008).

Currently in HTESSEL only grid-scale lakes (lake fraction larger than 50%) are considered. In these grid points, the surface temperature is prescribed, constant during integration, from the assimilation system of sea surface temperature for points close to large ocean bodies, and the remaining points are set to a monthly climatology. This treatment of lakes in the model does not account for the diurnal changes of surface temperature and inter-annual variability. The model does not consider partial lake cover within a grid-box. All the grid-points with lake fraction less then 50% are treated as land only. The current lake representation in the model has two main drawbacks: (i) the continuous increase in horizontal resolution of the forecast system leads to an increase in the number of resolved lakes with ill-defined properties, and (ii) sub-grid fraction occupied by lakes is not represented in coarser resolutions (used for seasonal or climate applications).

FLake is a numerical model, describing the thermodynamic processes in a lake, including the temperature of the water mass, underlying bottom sediments, ice and snow cover (Mironov 2008). Mironov *et al*. (1991) described a model that served as prototype to develop FLake and Mironov *et al*. (2010) presented a brief description of FLake, along with its implementation into a limited-area atmospheric model. FLake can be considered as a 0.5-dimensional model, since it has intermediate features between purely bulk models and multi-layer one-dimensional models. The vertical temperature profile specified in FLake consists of a top mixed layer, with uniform distribution of temperature, and a thermocline, with its upper boundary located at the mixed layer bottom, and the lower boundary at the lake bottom. The profile in the thermocline is parameterized according to the self-similarity concept (Kitaigorodski and Miropolski 1970), which has been supported by many theoretical (Barenblatt 1978, Zilitinkevich and Mironov 1992) and observational studies (Tamsalu and Myrberg 1998). In FLake, this concept is also applied to calculate temperatures in the snowpack, ice cover and the layer of bottom sediments. The above mentioned assumptions lead to a significant model simplification in comparison with multi-layer one-dimensional models, while still containing much more realistic physics than bulk models. In particular, the unrealistic assumption of fully mixed temperature profile in a lake, which has been applied in many bulk models, is avoided here. Instead, the mixed layer depth is calculated taking into account both wind and thermally driven mixing (Mironov 2008). The volumetric heating of the water column by solar radiation is also accounted for. Therefore, FLake meets the requirements for physical parameterizations in weather forecast systems, in terms of both physical realism and numerical efficiency. The FLake model package, available at http://lakemodel.net, contains the source code of the model itself and some auxiliary procedures, e.g. subroutines that calculate turbulent fluxes at the lake–atmosphere interface.

A new tile (ninth) was created in HTESSEL to represent the lakes interface with the atmosphere. Since turbulent fluxes are already calculated inside HTESSEL for each tile, FLake subroutines calculating these fluxes over a lake were omitted. Another simplification of the original FLake model deals with the snow cover and bottom sediments. HTESSEL has its own snowpack parameterization, which is used in this study for frozen lakes instead of the original FLake self-similarity model to be consistent with other parts of the code. HTESSEL snow scheme was coupled to the new lake tile, allowing the representation of snow cover over frozen lakes. The other simplification was to neglect bottom sediments. Considering the present global application for numerical weather prediction, the representation of bottom sediment would imply an extra set of external parameters which are not well known. In addition to the technical motivation, the heat storage and response time scales associated with bottom sediments in relatively large lakes (typically more than several meters deep) can be neglected when compared with the remaining energy forcings acting on a lake. As an extra physical motivation, one may notice that the parameterization of heat transfer in bottom sediments used in FLake is based on the self-similarity concept of the temperature profile (Mironov *et al*. 2003). This concept application has been extensively validated for the water column, but not to the bottom sediments.

Data

ERAI reanalysis provided the set of near-surface meteorological forcing to drive the LSM in stand-alone mode. ERAI covers the period of January 1989 to December 2005 (real time update is expected). The atmospheric forcing data were gridded on the original Gaussian reduced grid N128 (resolution of 0.7° over the equator) globally on a domain of 25 979 grid points (land and lake points only) at a three-hour interval. The state variables of the surface pressure, air temperature, specific humidity and wind were provided as instantaneous values from the lowest model level (approximately 10 m above the surface). State variables correspond to the 3–12-h forecast interval from initial conditions at 00:00 and 12:00 UTC. Surface fluxes represent three-hour averages and include precipitation, snowfall, and downwelling solar and thermal radiation. To avoid the initial spin-up in precipitation, the three-hourly surface fluxes are taken from the 9 h to 21 h forecasts initialized at 00:00 and 12:00 UTC.

The remotely-sensed lake surface water temperature (LSWT) derived from the observations taken by the Moderate-Resolution Imaging Spectroradiometer (MODIS) instrument aboard the NASA TERRA satellite platform in the original 4 km resolution (http://oceancolor.gsfc.nasa.gov/) were aggregated to the ERAI grid. The weekly composites from January 2001 to December 2005 of daytime and nighttime data were averaged, resulting in single weekly time series for each ERAI grid-box. This dataset has been mainly prepared for ocean sciences (Esaias *et al*. 1998), but it has also been applied to derive LSWT of the Great Salt Lake (Crosman and Horel 2009).

Observations of the lake ice duration were taken from the Global Lake and River Ice Phenology (Benson and Magnunson 2007). This dataset contains freezing and breakup dates and some other ice-cover characteristics for 750 lakes and rivers. Of the 429 water bodies that have records longer than 19 years, 287 are in North America and 141 are in Eurasia. Krinner (2003) used this dataset to validate lake simulations within a climate model, and Magnuson *et al*. (2000) analysed trends in lakes and rivers ice cover. Using the raw data, the mean lake ice duration was calculated by aggregating all the data within each grid-box of ERAI and calculating the mean from all the lakes and available periods. The lake characteristics of each lake within a specific grid-box might vary significantly. However, in the present modelling framework each grid-box can only have one lake that responses to the grid-box mean surface energy and water fluxes.

Simulations setup

Global offline simulations forced by ERAI were performed for the period January 1989 to December 2005. Since lake depth is an important parameter in the lake thermodynamics, three simulations were performed with globally constant lake depths of 10, 30 and 50 metres. A constant lake depth approach was chosen, since information on lake depths is not available globally. Other global studies of lakes within GCMs also used global uniform lake depths (e.g. Bonan 1995). One control experiment, with the lake model deactivated, was also performed. In this simulation, grid-points with a lake fraction less than 50% were treated as land and all the grid-points with a lake fraction larger than 50% (resolved lakes) were treated as water with time-constant temperature. To examine the effect of the snow insulation effect on lake ice, one extra set of simulations was performed with the lake ice snow-free. In the following discussion, WL (with lake, and overlaying snow deck) and WLns (with lake, but no snow on top of lake ice) refer to simulations with the lake model activated with snow deck and no snow deck on the lake ice, respectively, and NL (no lake) refers to simulations with the lake model switched off. The evolution of lake ice with no overlying snow (always the case in WLns simulations and when there is no snowfall in WL simulations), is mainly controlled by the albedo. In the present implementation, the lake ice albedo was equal to that of sea ice (already defined in HTESSEL) which is prescribed as a monthly climatology from Ebert and Curry (1993). Initial conditions were derived by performing one 11-year run (1989 through 1999) for each simulation, starting from arbitrary values. The first year of simulation was discarded and the following ten years were retained to create climatological initial conditions. This initialization methodology guarantees enough spin up time to stabilize the

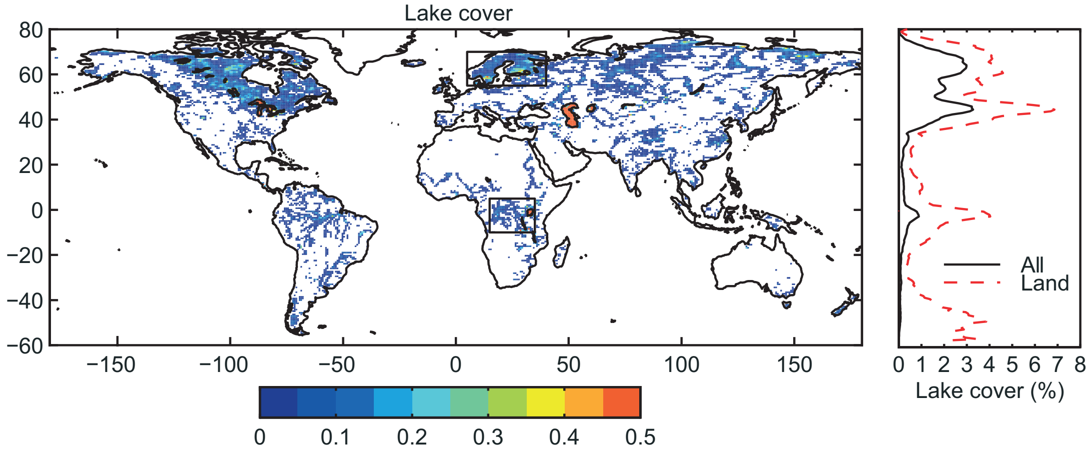

Fig. 1. Lake fraction in the Gaussian reduced grid N128 and zonally averaged fractional lake cover with respect to land points only (dashed) and all points (solid).

lake model results, in particular for simulations with lake depths of 50 meters.

The lake fraction was dervived from the GLCC land-cover dataset (Fig. 1; Loveland *et al*. 2000). It includes all inland open water surfaces, and does not distinguish between lakes and rivers. Besides the lake fraction, all the other input fields (e.g. land sea mask, vegetation type and cover, monthly albedos, soil types) were the same as in ERAI.

Results

Validation

An eight-day running mean was applied to the modelled LSWTs (*see* Fig. 2). This allowed for a coherent comparison with the weekly composites of remotely sensed LSWTs. For high-latitude lakes (Fig. 2a), the RMSEs varied significantly with the total lake depth. When exposed to the same meteorology, deeper lakes had higher heat content, leading to reduced ice cover durations than in shallower lakes. In general, deeper lakes also occupy larger areas (Lehner and Doll 2004), resulting in more intense lateral mixing/transport that is not represented by FLake. This affects the timing of freezing/melting.

For each grid point, the lowest LSWT RMSE was selected from one of the three simulations with different lake depths (BEST in Fig. 2). This simple approach is able to improve the LSWT RMSE distribution for high latitude lakes, when compared with globally constant lake depth simulations. These results confirm that a lake-depth tuning procedure, based on LSWT RMSE, would be able to improve the model skill in representing surface temperatures. In low-latitude lakes, the RMSEs did not vary significantly with the lake depth (Fig. 2b). Low-latitude lakes have normally a strong vertical stratification during the entire year due to a large amount of available solar energy. With this thermodynamic behaviour, the surface temperature is not as strongly constrained by the lake depth as it is for high-latitude lakes. The overall RMSEs were higher for high-latitude lakes. However, the ratio between RMSEs and the amplitude of the mean annual cycle of LSWT was higher in low-latitude lakes.

Further analysis was performed by removing the systematic differences between observed and modelled temperature before the RMSE calculation. After this procedure, only the BEST RMSE (lower RMSE in each grid point from different simulations) was chosen (BEST (bias removed) in Fig. 2). In this way, we distinguished the systematic errors from variability errors of modelled LSWTs. In high-latitude lakes, the systematic bias removal had a small impact on the statistical distribution. On the other hand, in low-latitude lakes most of the errors were associated with systematic biases. These differences between the partitioning of systematic and variability errors

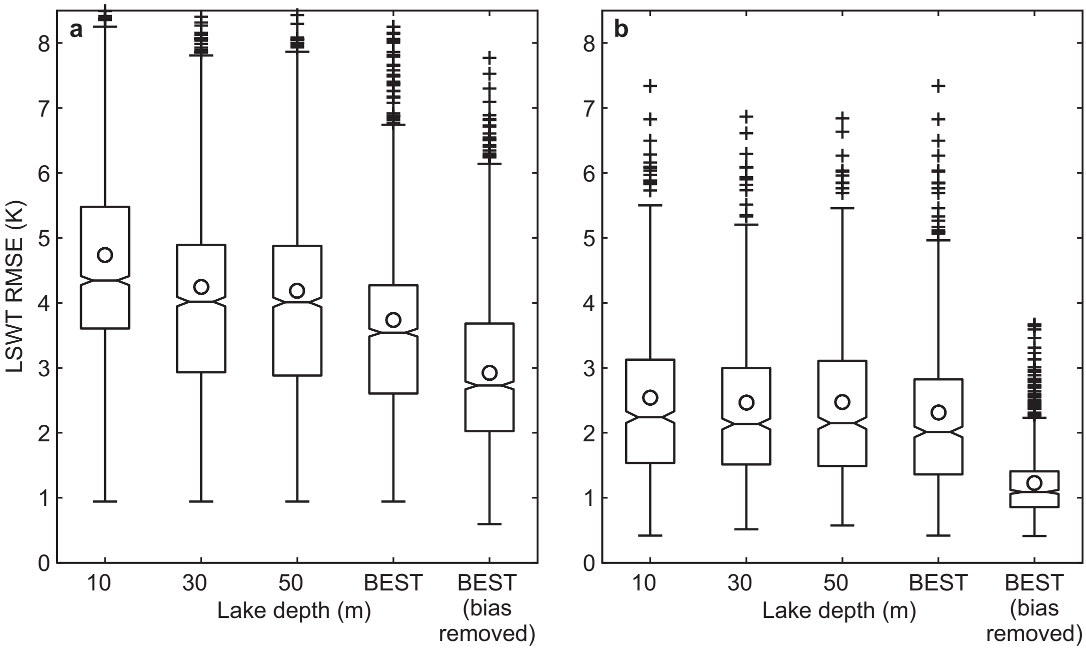

Fig. 2. Root mean square error distributions of the model *vs.* MODIS observations of lake surface-water temperature (LSWT, WL) of (**a**) high latitude (latitude > 40°) lakes, and (**b**) low latitude (latitude < 40°) lakes for all ERAI grid-boxes with more than 10% of available remotely sensed LSWTs data in the period January 2001 to December 2005. Separate statistics are given, since the dependence of the lake thermodynamics on the latitude is very pronounced due to the impact of freezing in high-latitude lakes. Distributions of RMSE are presented as function of lake depths (10, 30 and 50 m) and BEST (lake depth that minimizes the RMSE for each grid point). The rightmost "BEST (bias removed)" give results with the model-observations bias removed before computing RMSE. The boxes have horizontal lines at the lower quartile, median, and upper quartile, and outliers are data with values beyond 1.5 times the inter-quartile range indicated with '+'. Open circles are the means. Remotely-sensed LSWTs missing data are due to (i) cloud covered pixels, (ii) frozen lakes (the retrieval algorithm can only be applied for open water surfaces), and (iii) other errors in the retrieval algorithm.

for high and low latitude lakes can be explained as follows: (i) the errors in high-latitude LSWTs arise mainly from the inaccurate representation of the formation/disappearance of lake ice, and (ii) the errors in low-latitude lakes are mainly systematic. The systematic errors found can, in part, be associated with the forcing (e.g. uncorrected temperature due to differences between model orography and real lake altitude).

Simulations of lake ice duration were compared against Global Lake and River Ice Phenology database (Benson and Magnunson 2007). Model data were averaged for the entire simulation period (1989–2005) (Fig. 3). For each lake-depth simulation (including BEST), the distributions are displayed side by side for the WL and WLns (indicated with *). The WLns simulations were conducted in order to evaluate the importance of the snow insulation effect in the lake ice thermodynamics. The snow insulation effect reduces the coupling between the atmosphere and underlying surface (Cook *et al*. 2008). The simulations WLns produced higher positive errors in the lake-ice duration, as compared with the WL simulations. Snow insulates the lake ice, preventing its growth during the cold season, and thus reducing its duration. Despite the wide dispersion of errors in lake ice durations, 75% of the errors in the BEST lake depth selection were within ±30. These errors refer to the total lake ice duration. If the errors associated with the freezing and breakup were similar, 75% of the evaluated grid points would lie within an error margin of two weeks for the formation and breakup of lake ice. It should be noted that the time periods covered by the lake ice observations are not nec-

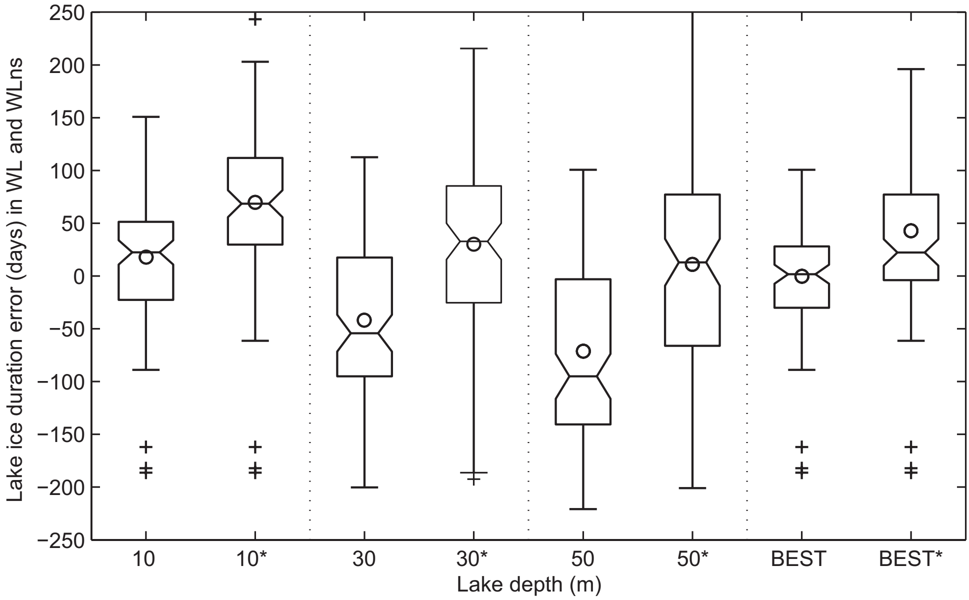

Fig. 3. Distributions of lake ice duration errors (observations minus model) from the simulations with lake (WL), and with lake and no snow (WLns). The latter is indicated with * in the horizontal axis. Data are from 123 grid points in the northern hemisphere. The boxes have horizontal lines at the lower quartile, median, and upper quartile, and outliers are dates with values beyond 1.5 times the inter-quartile range indicated with '+'. Open circles are the means.

essarily the same as the simulations. Additionally, the spatial averages (for all the lakes lying in each grid-box) can mix lakes that can have completely different error characteristics, leading to possible biases in the comparisons.

To further examine the impact of snow coupling over frozen lakes, lake ice durations, given by both coupled (WL) and uncoupled snow (WLns) simulations were performed (Fig. 4a). The lake ice duration was calculated for each year as the number of days with ice cover from September of the previous year to August of the following year. Simulated ice cover durations were always longer in WLns than in WL. These results reinforce the importance of the snow insulation effect on the thermodynamics of the underlying surface. Snow deposition over the lake ice reduces its cooling and growth. The formation period of lake ice was very similar in both WL and WLns simulations, while the breakup was delayed in WLns due to excessive ice thickness. As discussed before, the WLns simulations had a positive bias in lake ice duration (Fig. 3). These results show that coupling the snow deck with the lake ice improves the model skill to properly represent the lake ice melting. Lake ice albedo was prescribed as a climatology (Ebert and Curry 1993) in both WL and WLns simulations. In WL simulations this is not critical, since during most of the cold season the lake ice is covered by snow. On the other hand, the lake ice albedo controls the amount of solar radiation penetrating the ice in WLns. The Ebert and Curry (1993) albedo climatology might not be suitable for lake ice, which could explain the results of WLns simulations. A possible solution to mimic the snow insulating effect would be to change the lake ice albedo, but this is beyond the scope of the present work.

The effect of snow on frozen soil is similar to that of snow on ice: a snow-free soil will be colder and the soil melting will be delayed as compared with a snow-covered soil. In order to highlight the similarity of the effects, a comparison between the duration of frozen soil and lake ice was made (Fig. 4b, including only the grid points where lakes are represented). The frozen soil duration was evaluated by calculating the number of days, in each period from September to August, when the top soil layer (7-cm depth) was frozen. The

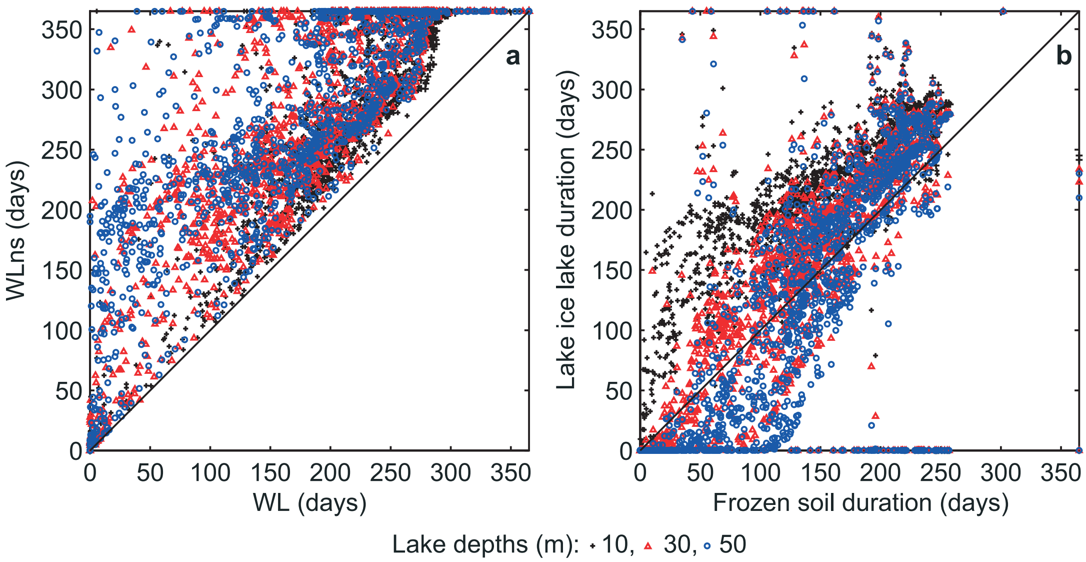

Fig. 4. (**a**) Mean lake ice duration simulations from WLns (with lake, no snow) *versus* the same quantity for WL (with lake) simulations. (**b**) Mean lake ice duration (only sub-grid scale lakes), in WL simulations, as a function of the mean frozen soil duration for the same grid-box in the 17-year simulation.

frozen-soil period durations of 150 to 250 days were similar to those for lake ice for all lake depths. For frozen soil durations of less than 150 days, the relation with lake ice duration strongly depended on the lake depth. Shallow lakes stayed frozen for longer periods, while deeper lakes might not even freeze. In the scatter plot (Fig. 4b), many symbols lying at or just above the horizontal axis, correspond to grid points where the soil froze but the lakes (in the same grid box) did not. This behaviour can mainly be attributed to the higher thermal inertia of lakes when compared with soil. It is reasonable to assume that in many locations, the top soil layer can freeze but nearby lakes remain unfrozen.

Impact on surface fluxes

In the following, we discuss only the simulations NL and WL with a globally constant lake depth of 30 m. Comparisons between the WL simulations with different lakes depths showed that on a monthly time scale the differences are not significant. The WLns simulations were also discarded from the following analysis due to the systematic biases on lake ice duration described before (*see* Fig. 3).

The main purpose of the implementation of the lake model within the LSM is to better represent the sub-grid cover variability and its impacts on land surface fluxes. Figure 5 shows the impact of sub-grid scale lakes on surface fluxes over two climatic regions: northern Europe (55°N–70°N, 5°E–40°E) and Africa (10°S–5°N, 15°E–35°E). These areas are represented as squares in Fig. 1 and comprise 385 and 585 grid-points for northern Europe and Africa, respectively. The analysis was performed for the mean annual cycle of the latent and sensible heat flux, solar and thermal net radiation, and surface net energy. The sign of the fluxes follow the model convention, with positive values for incoming and negative ones for outgoing energy from the surface. By using only grid points with a sub-grid lake cover > 5%, the mean annual cycles of area-averaged fluxes for the NL simulation (Fig. 5a and b) and for the difference WL – NL (Fig. 5c and d) were calculated. The differences were tested for significance with a two-tailed *t*-test. Grid points where sub-grid scale lake cover > 5%, composed 44% and 12% of the total analysed areas of northern Europe and Africa, respectively. Over northern Europe, there was a significant increase of the energy storage on the surface during summer and release during autumn. This kind of behaviour

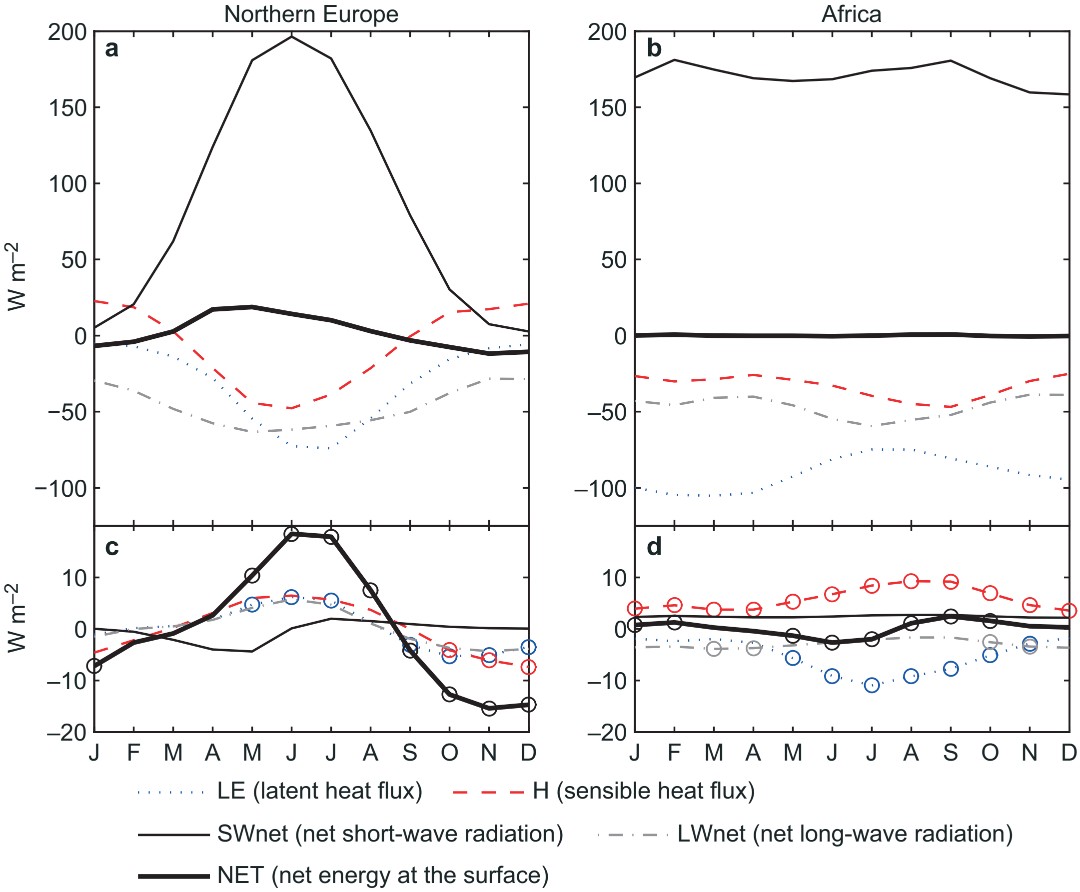

Fig. 5. Mean annual cycle from simulation WL (lake depth of 30 m) simulation of surface fluxes for all sub-grid scale lake grid points inside (**a** and **c**) northern Europe and (**b** and **d**) Africa (*see* Fig. 1 for areas definition). **c** and **d**: differences between simulations with lake minus no lake (WL – NL); circles mark differences significant at $p < 0.01$ (two-tailed t-test).

has been already documented in other works (e.g. Bonan, 1995), with cooling and warming effects on the atmosphere during summer and autumn, respectively. Over Africa, the impacts of the representation of sub-grid scale lakes differed from those at high latitudes. The net surface energy remained similar throughout the year, but there was a significant change in the flux partitioning, indicating an increase of the latent heat flux and decrease of the sensible heat flux. This decrease in the Bowen ratio (ratio between sensible and latent heat fluxes) can impact the boundary layer development and triggering convection. Over northern Europe, there was a reduction of the net short-wave radiation during spring due to an increase of the surface albedo, but the differences were not significant. In both areas the changes in thermal net radiation were small and not significant in most of the months. This analysis was performed in other high and low latitude regions with similar results.

A global overview of the impact on evaporation is presented in Fig. 6. The setup of the NL simulation included constant LSWT during the run, which is not realistic. To overcome this problem, evaporation for all resolved lakes of NL simulation was replaced by ERAI data. The mean annual evapotranspiration from the NL simulation (Fig. 6a) was compared against ERAI (not shown). The main differences between NL and ERAI evapotranspiration can be primarily attributed to the soil moisture assimilation scheme used in ERAI. The assimilation corrects

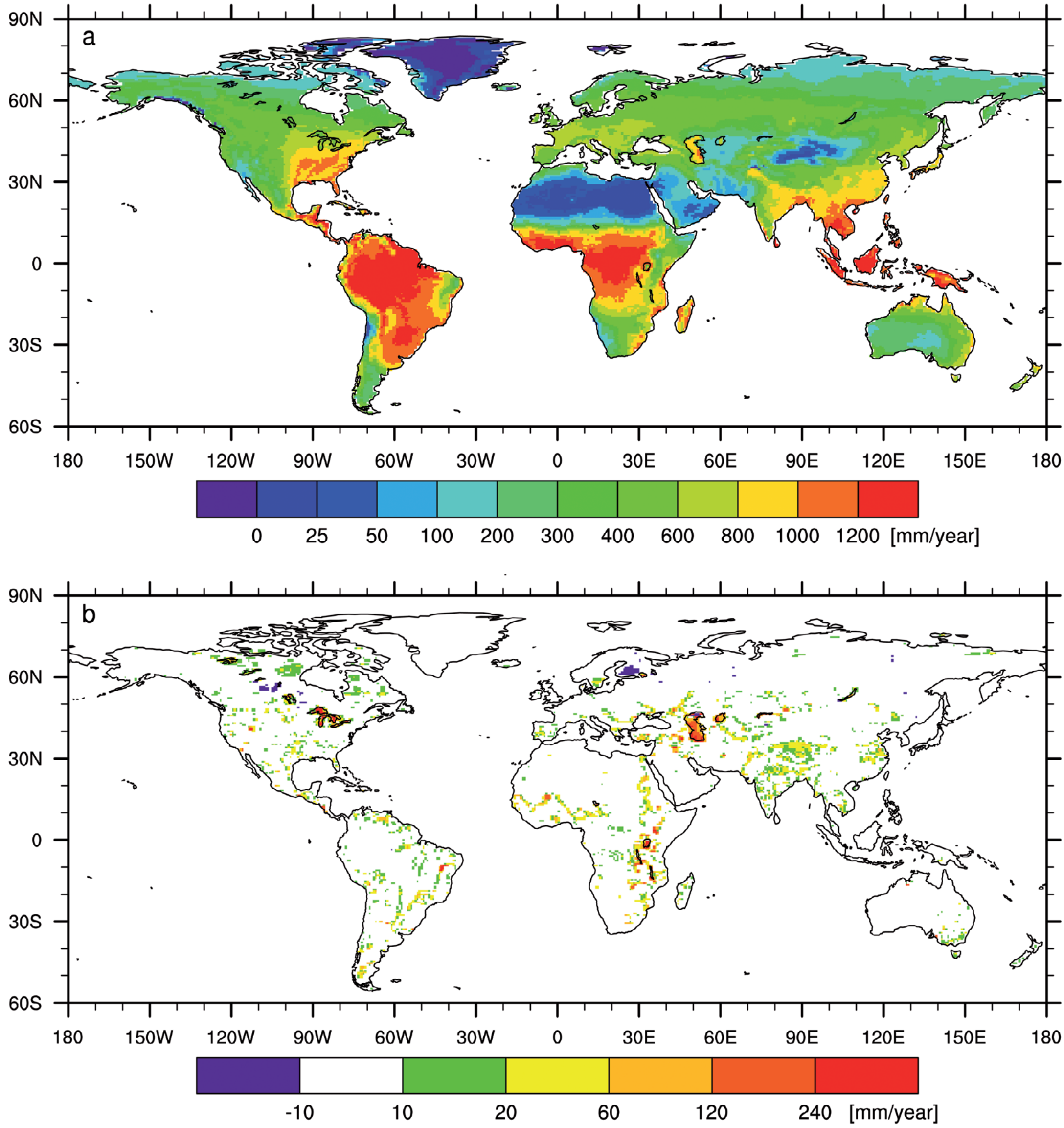

Fig. 6. (**a**) Global yearly evapotranspiration from the simulation NL (no lakes). (**b**) Yearly evapotranspiration differences between the simulations WL (lake depth 30 m) and NL (no lakes). Evaporation for all resolved lakes of NL simulation was replaced by ERAI data. Note the different colour scales.

soil moisture in the first three layers based on 6-hour increments of atmospheric analysis of specific humidity and temperature at the lowest model level (Douville *et al*. 2000). When ERAI soil moisture increments were added to the evaporation, a good match with NL was found (not shown). A good agreement between offline and ERAI evapotranspiration shows that the standalone modelling methodology is able to reproduce, apart from the assimilation, the evapotranspiration fluxes of the reanalysis. The patterns of impact of the representation of lakes (both grid and sub-grid scale) on evaporation (Fig. 6b) were very similar to those of the lake cover fraction (Fig. 1), with an increase of evaporation throughout the globe. On a yearly scale, the differences were small when compared with the total evapotranspiration, and very localized. Nevertheless, the impact can be significant in some specific areas and seasons.

Summary and discussion

This work reports the implementation of the FLake lake model into the LSM HTESSEL. Stand-alone global simulations forced by the ERAI reanalysis were performed and LSWT was validated against remotely sensed LSWTs derived from the MODIS TERRA satellite. Total lake depth appears to be the most important model parameter, but global information on lake depths is not available. To overcome this lack of information and to examine the model sensitivity to lake depths, simulations were performed with three different lake depths of 10, 30 and 50 meters. A simple minimizing strategy of choosing the lake depth (for each grid point) with the lowest RMSE was able to reduce significantly the LSWT RMSEs distribution. This tuning exercise proves to be effective for obtaining a first guess for a global-wide lake depth dataset (for all lakes with available remotely sensed LSWT). Further validation was performed by comparing simulated and observed lake ice duration. Regarding the lake ice evolution, the coupling with the current snow scheme of HTESSEL showed significant improvements. The snow insulation effect was analysed in terms of the lake ice duration and compared against frozen soil at nearby locations.

The main drawback of the presented stand-alone simulation methodology is related to the fixed atmospheric forcing, with no representation of the atmospheric response to the presence of lakes. It is expected that adding a new lake component to the land surface will change the near-surface atmospheric fields. Since this new component will evaporate at the potential rate, one should assume that especially the near-surface relative humidity will change. A regional study based on measurements from high-latitude lakes (Rouse *et al*. 2005) found that lakes strongly enhanced evapotranspiration when added to a landscape, especially during autumn when vertical temperature and humidity gradients are larger. With more abundant water supply to the atmosphere, an increase of the relative humidity is expected. However, a stronger free convective regime can also enhance a deepening of the boundary layer, thus reducing the relative humidity at the surface. An increase (or decrease) of the screen-level relative humidity will modify the water demand from the atmosphere, but such equilibrium can only be tested with coupled simulations.

In spite of the limitations of the stand-alone simulation methodology, the model performs as expected when compared with observed LSWTs and lake-ice duration. The impact of sub-grid scale lakes on surface fluxes was analysed on a monthly scale and in different climates. In high-latitude regions, lakes change the terrestrial energy storage cycle, with increased storage during the summer and subsequent release during the autumn. In equatorial regions, lakes change the partition of energy fluxes to the atmosphere, with an increase of the latent heat flux and reduction of the sensible heat flux.

Further simulations are needed to properly investigate the impact of lakes on surface fluxes. Coupled atmosphere simulations are necessary but require significant computational resources. Nevertheless, the present work shows that lakes significantly impact on the surface energy and water balance. Global lake depths are required, it is the most important lake parameter. Besides the lake depth, the lack of global data on other lake related parameters, such as optical characteristics of lake water, is still an impediment to further progress.

Acknowledgements: This work was supported by the Portuguese Foundation for Science and Technology (FCT) under project AMIC PTDC/AAC-CLI/109030/2008 cofinanced by the European Union under program FEDER. E. Dutra acknowledges the financial support of FCT under the grant SFRH/BD/35789/2007, and expresses his gratitude to ECMWF for travel support. Viktor Stepanenko's stay in Lisbon was supported by the European Commission FP6 Integrated Project WATCH, proj. no. 036946.

References

Balsamo G., Viterbo P., Beljaars A., Van den Hurk B., Betts A.K. & Scipal K. 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.

Barenblatt G.I. [Баренблатт Г.И.] 1978. [Self-similar distribution of temperature and salinity in upper thermocline]. *Izvestiya Akademii Nauk SSSR, Fizika Atmosfery i Okeana* 14: 1160–1166. [In Russian].

Benson B. & Magnunson J. 2007. *Global lake and river ice* *phenology database*. CD-ROM, National Snow and Ice Data Center/World Data Center for Glaciology, Boulder, Colorado.

Bonan G.B. 1995. Sensitivity of a GCM simulation to inclusion of inland water surfaces. *J. Climate*. 8: 2691–2704.

Cook B.I., Bonan G.B., Levis S. & Epstein H.E. 2008. The thermoinsulation effect of snow cover within a climate model. *Clim. Dyn.* 31: 107–124.

Crosman E.T. & Horel J.D. 2009. MODIS-derived surface temperature of the Great Salt Lake. *Remote Sens. Environ.* 113: 73–81.

Dirmeyer P.A., Dolman A.J. & Sato N. 1999. The pilot phase of the Global Soil Wetness project. *Bull. Am. Meteorol. Soc.* 80: 851–878.

Douville H., Viterbo P., Mahfouf J.F. & Beljaars A.C.M. 2000. Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data. *Mon. Weather Rev.* 128: 1733–1756.

Dutra E., Viterbo P. & Miranda P.M.A. 2008. ERA-40 reanalysis hydrological applications in the characterization of regional drought. *Geophys. Res. Lett.* 35, L19402, doi:10.1029/2008GL035381.

Ebert E.E. & Curry J.A. 1993. An intermediate one-dimensional thermodynamic sea-ice model for investigating ice–atmosphere interactions. *J. Geophys. Res.* 98: 10085–10109.

Esaias W.E., Abbott M.R., Barton I., Brown O.B., Campbell J.W., Carder K.L., Clark D.K., Evans R.H., Hoge F.E., Gordon H.R., Balch W.M., Letelier R. & Minnett P.J. 1998. An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.* 36: 1250–1265.

Jarlan L., Balsamo G., Lafont S., Beljaars A., Calvet J.C. & Mougin E. 2008. Analysis of leaf area index in the ECMWF land surface model and impact on latent heat and carbon fluxes: application to West Africa. *J. Geophys. Res.* 113, D24117, doi:10.1029/2007JD009370.

Hostetler S.W., Giorgi F., Bates G.T. & Bartlein P.J. 1994. Lake–atmosphere feedbacks associated with paleolakes Bonneville and Lahontan. *Science* 263: 665–668.

Kitaigorodski S.A. & Miropolski Yu.Z. [Китаигородски С.А. & Миропольски Ю.З.] 1970. [On the theory of open ocean active layer]. *Izvestiya Akademii Nauk SSSR, Fizika Atmosfery i Okeana* 6: 178–188. [In Russian].

Krinner G. 2003. Impact of lakes and wetlands on boreal climate. *J. Geophys. Res.* 108(D16), 4520, doi:10.1029/2002JD002597.

Lehner B. & Doll P. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* 296: 1–22.

Lofgren B.M. 1997. Simulated effects of idealized Laurentian Great Lakes on regional and large-scale climate. *J. Climate* 10: 2847–2858.

Long Z., Perrie W., Gyakum J., Caya D. & Laprise R. 2007. Northern lake impacts on local seasonal climate. *J. Hydrometeor.* 8: 881–896.

Loveland T.R., Reed B.C., Brown J.F., Ohlen D.O., Zhu Z., Yang L. & Merchant J.W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* 21: 1303–1330.

Magnuson J.J., Robertson D.M., Benson B.J., Wynne R.M., Livingstone D.M., Arai T., Assel R.A., Barry R.G., Card B., Kuusisto E., Granin N.G., Prowse T.D., Stewart K.M. & Vuglinski V.S. 2000. Historical trends in lake and river ice cover in the northern hemisphere. *Science* 289: 1743–1746.

Mironov D. 2008. *Parameterization of lakes in numerical weather prediction. Description of a lake model.* COSMO Technical Report no. 11, Deutscher Wetterdienst, Offenbach am Main, Germany.

Mironov D., Golosov S.D., Zilitinkevich S.S., Kreiman K.D. & Terzhevik A.Yu. 1991. Seasonal changes of temperature and mixing conditions in a lake. In: Zilitinkevich S.S. (ed.), *Modelling air–lake interaction: physical background*, Springer-Verlag, Berlin, pp. 74–90.

Mironov D.V., Golosov S.D. & Zverev I.S. 2003. Temperature profile in lake bottom sediments: an analytical self-similar solution. In: Terzhevik A.Yu. (ed.), *Proceedings of the 7th Workshop on Physical Processes in Natural Waters*, Northern Water Problems Institute, Russian Academy of Sciences, Petrozavodsk, Karelia, Russia, pp. 90–97.

Mironov D.V., Heise E., Kourzeneva E., Ritter B., Schneider N. & Terzhevik A. 2010. Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Env. Res.* 15: 218–230.

Rouse W.R., Oswald C.J., Binyamin J., Spence C.R., Schertzer W.M., Blanken P.D., Bussieres N. & Duguay C.R. 2005. The role of northern lakes in a regional energy balance. *J. Hydrometeor.* 6: 291–305.

Rutter N., Essery R., Pomeroy J., Altimir N., Andreadis K., Baker I., Barr A., Bartlett P., Boone A., Deng H.P., Douville H., Dutra E., Elder K., Ellis C., Feng X., Gelfan A., Goodbody A., Gusev Y., Gustafsson D., Hellstrom R., Hirabayashi Y., Hirota T., Jonas T., Koren V., Kuragina A., Lettenmaier D., Li W.P., Luce C., Martin E., Nasonova O., Pumpanen J., Pyles R.D., Samuelsson P., Sandells M., Schadler G., Shmakin A., Smirnova T.G., Stahli M., Stockli R., Strasser U., Su H., Suzuki K., Takata K., Tanaka K., Thompson E., Vesala T., Viterbo P., Wiltshire A., Xia K., Xue Y.K. & Yamazaki T. 2009. Evaluation of forest snow processes models (SnowMIP2). *J. Geophys. Res.* 114, D06111, doi:10.1029/2008JD011063.

Simmons A.J., Uppala S.M., Dee D. & Kobayashi S. 2007. ERA-Interim: new ECMWF reanalysis product from 1989 onwards. *ECMWF Newsletter* 110: 25–35.

Tamsalu R. & Myrberg K. 1998. A theoretical and experimental study of the self-similarity concept. *MERI, Report Series of the Finnish Institute of Marine Research* 37: 3–13.

Uppala S.M., Kallberg P.W., Simmons A.J., Andrae U., Bechtold V.D., Fiorino M., Gibson J.K., Haseler J., Hernandez A., Kelly G.A., Li X., Onogi K., Saarinen S., Sokka N., Allan R.P., Andersson E., Arpe K., Balmaseda M.A., Beljaars A.C.M., Van De Berg L., Bidlot J., Bormann N., Caires S., Chevallier F., Dethof A., Dragosavac M., Fisher M., Fuentes M., Hagemann S., Holm E., Hoskins

B.J., Isaksen L., Janssen P.A.E.M., Jenne R., McNally A.P., Mahfouf J.F., Morcrette J.J., Rayner N.A., Saunders R.W., Simon P., Sterl A., Trenberth K.E., Untch A., Vasiljevic D., Viterbo P. & Woollen J. 2005. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* 131: 2961–3012.

Van den Hurk B. & Viterbo P. 2003. The Torne-Kalix PILPS 2(e) experiment as a test bed for modifications to the ECMWF land surface scheme. *Global Planet. Change* 38: 165–173.

Van den Hurk B., Viterbo P., Beljaars, A.C.M. & Betts A.K. 2000. Offline validation of the ERA40 surface scheme. *ECMWF Tech. Memo.* 295: 1–42.

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

Zilitinkevich S.S. & Mironov D.V. 1992. Theoretical model of the thermocline in a freshwater basin. *J. Phys. Oceanogr.* 22: 988–996.