Modeling subgrid lake energy balance in ORCHIDEE terrestrial scheme using the FLake lake model

Anthony Bernus and Catherine Ottlé

Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace (IPSL), CEA-CNRS-Université Paris-Saclay, Orme des Merisiers, 91190 Gif-sur-Yvette, France

Correspondence: Anthony Bernus and Catherine Ottlé (anthony.bernus@lsce.ipsl.fr ; catherine.ottle@lsce.ipsl.fr)

Abstract.

The freshwater 1-D FLake lake model was coupled to the ORCHIDEE land surface model to simulate lake energy balance at the global scale. A multi-tile approach has been chosen to allow the modelling of various types of lakes within the ORCHIDEE grid cell. Thus, three different lake tiles have been defined according to lake depth which is the most influential parameter of FLake, but other properties could be considered in the future. Several depth parameterization strategies have been compared, differing by the way to aggregate the depth of the subgrid lakes, i.e., arithmetical, geometrical, harmonical mean and median. Five atmospheric reanalysis datasets available at $0.5°$ or $0.25°$ resolution, have been used to force the model and assess model systematic errors. Simulations have been performed, evaluated and intercompared against observations of lake water surface temperatures provided by the GloboLakes database over about 1000 lakes and ice phenology derived from the Global Lake and River Ice Phenology database.

The results highlighted the large impact of the atmospheric forcing on the lake energy budget simulations and the improvements brought by the highest resolution products (ERA5 and E2OFD). The median of the Root Square Mean Errors (RMSE) calculated at global scale, ranges between $3.2\ °C$ and $2.7\ °C$ among the forcings, CRUJRA and ERA5 leading respectively to the worst and best results. Depth parameterization strategy appeared to be less influent, with RMSE differences less than 0.1 $°C$ for the four aggregation scenarios tested.

The simulation of ice phenology presented systematic errors whatever the forcing and the depth parameterization used. Large systematic errors were highlighted such as negative biases on the onset and positive biases on the offset. Freezing onset was shown to be the less sensitive to atmospheric forcing with the median of the errors ranging between 10 and 14 days. Larger errors up to 25 days were observed on the simulation of the end of the freezing period. Such errors already highlighted in previous works, could be explained by scale effects and deficiencies in the modeling of snow/ice processes not accounting for partial ice cover. Various pathways are drawn to improve the model results, including the use of remote sensing data to better constrain the lake radiative parameters (albedo and extinction coefficient) as well as the lake depth thanks to the recent and forthcoming high-resolution satellite missions.

1 Introduction

25 Lakes are an important component of the continental hydrology and interfere in the global carbon cycle. Lakes do play a role in many physical and biogeochemical processes at the land-atmosphere interface such as the transfers of energy, water, carbon, and other greenhouse gases (GHGs) like methane. However, since they represent only a small percentage of the land surface (about 3.7% as shown by Verpoorter et al. (2014)), their contribution at the global scale is small (at least regarding the energy and water budgets). Therefore, lakes started to be explicitly represented in Earth System Models (ESMs) only recently, thanks

30 to the increased grid resolutions and computer advances.

Lake distribution is spatially unequal all around the world with two regions in the high latitudes (Canada and Scandinavia) showing the highest number of lakes (Downing et al., 2006). The second characteristic of this distribution is the presence of large lakes in Northern America and Africa which represent 27% of the world's total lake water volume and 11% of the lake surface (Messager et al., 2016). Their role on the physical properties of the low atmosphere and the importance to model these

35 processes, has been demonstrated by many studies. For example, Bonan (1995) showed that the implementation of lakes in the NCAR-CCM2 climate model leads to an increase of the latent heat fluxes transferred to the atmosphere and to a decrease of the annual amplitude of the surface temperature. They found that in the high inland lake regions in July, the average lake surface temperature was cooler by $2°C$ to $3°C$ and that the latent heat fluxes increased by 10 to 45 W/m^2 depending on the region.

Lakes also play a buffer role in the hydrological cycle impacting substantially the river regimes. As an example, in Alaska,

40 Bowling and Lettenmaier (2010) explain the reduction of the spring peak flow of the Putuligayuk river by the lake storage which represents up to 80% of the snowmelt.

Regarding the carbon and nitrogen cycles, lakes emissions are poorly constrained but it is recognized that their contribution is significant (Bastviken et al., 2004). Lakes are a carbon source for the atmosphere, with emissions of about 110 TgC of CO_2 in annual mean, but they are also a sink since an important part of organic carbon is buried into the bottom sediments

45 (Cole et al., 2007). Another important GHG is the methane generated from methanogesis by micro organisms in anaerobic conditions, processes which are presently poorly understood but could lead to lake emissions between 8 and $48 TgC-CH_4 y^{-1}$, comparable to the emissions of the oceans (Bastviken et al., 2004). For example, West et al. (2016) estimated that, between 8% and 16% of the non-anthropic methane emissions are coming from lakes. Recently, Beaulieu et al. (2019) gave a higher estimation of lake methane emissions with an amount of $112 Tg\, C - CH_4\, y^{-1}$.

50 Given the role of lakes on the global carbon and water cycles, the large uncertainties related to GHG budgets (Saunois et al., 2020) and the increasing spatial resolution of ESMs allowing to better account for the land surface diversity, it becomes necessary to implement water bodies in ESMs. At the scale of the ESMs (typically a few hundred of kilometers), lakes can not be explicitly represented and have to be aggregated in water fractions with effective parameters, allowing to solve the surface energy balance and to predict the time evolution of the surface temperature and fluxes exchanged with the atmosphere. Among

55 all the models developed for that purpose, the one-dimensional approaches are the most used in ESMs because of their low computational cost and the robustness of the parameterizations based on a few number of free parameters. In this category, the FLake model (Mironov, 2008) and the Hostleter model (Hostetler and Bartlein, 1990) are the most used to represent lakes

in ESMs. These two models follow different approaches to solve the lake energy budget and to calculate the energy vertical transfers and the water temperature profile: in FLake, the temperature profile is computed through the self-similarity theory whereas in Hostleter's model, the eddy diffusive equations are explicitly resolved. An intermediate methodology resolving the turbulence in the mixing layer with a bulk approach, was used by MacKay (2012) in the Canadian Small Lake Model (CSLM). This model was successfully implemented in the Environment and Climate Change Canada's surface prediction system (Garnaud et al., 2022). All these methods have shown good performances when compared at site level and participated to various intercomparison exercices. For example, in LakeMIP, Stepanenko et al. (2010, 2013) showed the good performances of a number of these models in predicting the water surface temperature and fluxes of various types of lakes.

As such, FLake has been implemented in numerous climate and numerical weather prediction models. For example, it has been coupled to the German regional atmospheric model (COSMO) and its TERRA land component (Mironov et al., 2010), the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric model and its HTESSEL land surface component (Balsamo et al., 2012; Dutra et al., 2010) or the French CNRM-CM5 ESM and its SURFEX interface (Le Moigne et al., 2016). Flake is also part of the JULES Land simulator (Rooney and Bornemann, 2013) of the UK Met Office Unified Model and can be run in a stand-alone version. The coupling of FLake with all these LSMs showed strong benefits in the climate and weather simulations. In addition, the LISSS (Subin et al., 2012) model based on the Hostetler model is used in the CLM land component of the CESM climate model developed by the NCAR institute (Oleson et al., 2013) and in the land models BCC-AVIM (Golaz et al., 2019) and NorESM2 (Seland et al., 2020). The LISSS model is also implemented in the LM4 land component of the Geophysical Fluid Dynamics Laboratory's ESM (Milly et al., 2014) and in the land component CLASS of the Canadian ESM (Martynov et al., 2012). More recently, mass balance equations have been implemented in some ESMs in order to solve the coupled energy and water cycles. For example, Huziy and Sushama (2017) implemented a lake mass budget in the CLASS LSM and tested it for regional applications. Besides, Guinaldo et al. (2021) developed the MLake mass budget model in SURFEX environment, thus enabling to improve the global scale water discharges and the seasonal evolution of lakes water levels.

This study aims to include lakes in the ESM of the Institut Pierre Simon Laplace (IPSL) and more precisely in its land surface component ORCHIDEE. Contrary to Krinner (2003) who directly introduced the Hostetler model in its atmospheric model component (LMDZ) and by this way, highlighted the role of water surfaces on the boreal climate, we chose here to represent explicitly the lakes in ORCHIDEE and to resolve, as a first step, the energy balance. For that purpose, the FLake lake model was coupled to ORCHIDEE and various developments were done to represent lakes in the grid cell, define effective parameters and set up the new parameterizations. We present these developments here and the results obtained through a series of simulations devoted to assess the respective role of the atmospheric forcing and of depth parameterization in the model performances. We focus on the analysis of the simulated surface variables (temperature, freezing state and turbulent fluxes) which are impacting the atmospheric surface boundary layer. Model performances are analyzed at the lake scale and globally, against satellite datasets. The contribution of lakes in the simulation of the surface temperature and energy fluxes is quantified. The following steps including further model developments and parameter optimization are finally discussed.

2 Data and Modeling framework

2.1 Datasets used

2.1.1 Land Cover map

The Land Cover CCI product (v2.0.7, Bontemps et al. (2015), ESA (2017)) developed in the framework of the Climate Change Initiative program of the European Space Agency (ESA) is used as a standard in ORCHIDEE to map the natural vegetation and derive the Plant Functional Types (PFTs) describing the subgrid variability of the vegetation. The product is available at a resolution of 300 m and the global maps are provided on a yearly basis, over the period (1992 – 2020). The interpretation of the 38 land cover classes in terms of ORCHIDEE PFTs is fully presented in Poulter et al. (2015) and Hartley et al. (2017), and detailed in the dedicated website https://orchidas.lsce.ipsl.fr/dev/lccci. In this work, we used the PFT maps generated at the resolution of $0.25°$ and $0.5°$ for the standard version of ORCHIDEE.

2.1.2 Lake area and depth

The HydroLAKES database (Messager et al., 2016) has been used to characterize lake fractions and depths. The dataset provides georeferenced polygons for 1.4 million lakes of size larger than $0,1\ km^2$. A mean depth is provided for each water body, estimated either from in situ measurements or from geomorphological considerations allowing to assess lake shape from which depth is derived. Additional attributes include estimates of the shoreline length, water volume and residence time but they were not used in this work. Moreover, all lakes are co-registered to the global river network of the HydroSHEDS database used in ORCHIDEE to describe the river system and constrain the water routing scheme. This specificity will then insure the coherence for the next developments around the modeling of the water budgets and the lakes-river interactions. The data can be downloaded at https://hydrosheds.org/page/hydrolakes.

2.1.3 Lake surface temperatures

The Global Observatory of Lake Responses to Environmental Change (GloboLakes) database is used to evaluate our simulations. This dataset has been developed thanks to a project funded by the UK Natural Environment Research Council in the continuity of the ARC-Lake database (MacCallum and Merchant, 2012) funded by the European Space Agency. It provides daily observations of Lake Surface Water Temperature (LSWT) for about 1000 lakes distributed all around the world, with uncertainties and quality level. The observations cover the period 1995-2016 with a spatial resolution of $0.05°$ and combine different instruments such as the AVHRR (Advanced Very High Resolution Radiometer) on MetOpA, AATSR (Advanced Along Track Scanning Radiometer) on Envisat and ATSR-1/2 (Along Track Scanning Radiometer) on ERS-1/2 (European Remote Sensing Satellite) covering different periods. The time series from ATSR-2 are available for the period 1995-2002, from AATSR for the period 2002-2010 and from AVHRR for the period 2007-2016. A common processing methodology and the fact that all the satellites have a quasi polar orbit leading to a constant local solar time, allowed a time homogeneity in the measurements and a homogeneous resulting dataset (Carrea and Merchant, 2019). Given the lack and lower quality of the data before 2002 due to

the older instruments, we used the reduced period 2000-2016 for the model evaluation. For the whole period between 1995 and

2009 and for the previous version of the database ARC-Lake, the product uncertainty estimated by Carrea and Merchant (2019)

125   is less than 1 $°C$. The comparison against local measurements show that errors are slightly larger for daytime measurements

(0.7 $°C$) compared to nighttime (0.5 $°C$) with a small negative bias (-0.2$°C$ during daytime and -0.1$°C$ at night) linked to the

thermal skin cooling effect. The data can be found at: http://dx.doi.org/10.5285/76a29c5b55204b66a40308fc2ba9cdb3.

2.1.4   Lake ice phenology

The Global Lake and River Ice Phenology Database is used to validate the lake freezing processes. This dataset provides

130   the start, end and duration of the freezing period for 857 lakes located in the northern latitudes. The dataset is based on in

situ measurements which are sometimes discontinuous and acquired over different periods, and reports the dates of full ice

coverage. The data are available for some lakes back to the year 1850 updated to 2020 (Benson, 2002). The data can be

downloaded at https://doi.org/10.7265/N5W66HP8 .

2.1.5   Atmospheric forcings

135   Different atmospheric reanalysis were used in this work to force the ORCHIDEE model and assess the model sensitivity to

meteorological forcing uncertainties. These reanalysis differ by the model and data used to constrain the atmospheric fields,

differences that may introduce some biases. Here, we used five different datasets available at three different spatial resolutions:

31 km, 0.25° and 0.5° and three different temporal resolutions, hourly, 3-hourly and 6-hourly.

Then, for the simulations performed at 0.5° resolution, we used: a) WFDEI-CRU, a 3-hourly 0.5° global forcing product

140   which is a combination of the ERA-interim reanalysis and the Climate Research Unit (CRU)-TS 3.1 climatological data (Harris

et al., 2014) presented by Weedon et al. (2011), b) CRUJRA, a 6-hourly 0.5° global forcing product, which is based on JRA-55

reanalysis (Kobayashi et al., 2015) corrected by CRU-TS 3.2 (Harris et al., 2014), and finally c) CRUNCEP, a 6-hourly 0.5°

global forcing product, which is a combination of two datasets: the CRU-TS 3.2 and the National Center for Environmental

Prediction (NCEP) reanalysis product described by Viovy (2018). ERA5 (Hersbach et al., 2020), a hourly global forcing

145   product at a resolution of 31 km, and E2OFD (Beck et al., 2017), a 3-hourly 0.25° global forcing product complete the selection.

All the data have been formatted by the ORCHIDEE team to be used and interpolated at the half-hourly time step, following

the methodology presented in Wei et al. (2014) and summarized in: https://forge.ipsl.jussieu.fr/orchidee/wiki/Documentation/

Forcings. In brief, the interpolation of the meteorological variables is done linearly except for the solar radiation where the

interpolation between two time step of the atmospheric forcing, follows a typical sinusoidal daily evolution according to the

150   solar zenithal angle. It should be noted that no extra atmospheric fields are needed to solve the energy balance at the lake –

atmosphere interface compared to what is done for the vegetated fraction of the grid.

2.2 Modeling framework

2.2.1 ORCHIDEE land surface model

ORCHIDEE is the land surface component of the IPSL ESM. It simulates the surface energy balance and the water, carbon
and nitrogen cycles for large scale and climate applications. The model is composed of two main modules named SECHIBA
and STOMATE. SECHIBA resolves the thermal and hydrological processes at a half-hourly time step, and computes the key
surface variables and fluxes such as the soil temperatures (Wang et al., 2016) and water content (de Rosnay, 2003) with a
11-layer discretization, the evolution of the snow cover with an explicit 3-layer representation (Wang et al., 2013) and the
photosynthesis of the vegetation (Vuichard et al., 2019). The second module STOMATE computes with a daily time step,
the fully coupled carbon and nitrogen dynamics and therefore calculates key processes such as the vegetation respiration, the
soil carbon dynamics, the litter decomposition, the vegetation phenology, constrained by the nutrients availability (Vuichard
et al., 2019). The model can be run in forced mode with atmospheric variables prescribed or coupled with the atmosphere
simulated by the atmospheric general circulation model LMDZ (Cheruy et al., 2020). In ORCHIDEE, the land surface is
represented by two Surface Functional Types (SFTs) differentiating bare and vegetated surfaces from glaciers. The first SFT
allows to represent the land surface into fractions of 15 different broad PFTs merging both plant species and main climate
zones, one of them representing bare soils. In the standard version of ORCHIDEE, urban areas and water surfaces are not
represented explicitly and are considered as bare soils. The fractions of PFTs assigned in a grid cell are determined by a land
cover map and a cross-walking-table (CWT) allowing to link the land cover classes to the ORCHIDEE PFTs. This procedure
is fully presented in Lurton et al. (2020). The model can be used at various scales, from the local to the global scale, with
grid resolutions constrained by the scale of the atmospheric input fields (incoming shortwave and longwave radiations, surface
air temperature and humidity, wind speed and precipitation). In this work, the model version used is the AR6 version which
participated to the CMIP6 model intercomparison exercice, and contributed to the IPCC 6th Assessment Report (Boucher et al.,
2020; Cheruy et al., 2020).

2.2.2 FLake lake model

FLake (Mironov, 2008) is a 1-D thermodynamic lake model primarily developed for weather forecast modeling purposes. It is
a bulk model capable of predicting the vertical temperature structure and mixing conditions in lakes, given the meteorological
conditions at the atmosphere interface (incoming radiation, air temperature and humidity, wind speed, snowfall). The water
temperature profile is represented by a single mixed layer with a uniform temperature above a thermocline. Extra modules are
implemented to model the snow, ice and sediment profile temperatures into specific layers.

The structure of the lake thermocline is parameterised using the concept of self-similarity. Therefore, the depth-temperature
relationship depends on a shape factor, which is resolved at each time step according to the boundary conditions. The same
approach of "assumed shape" is used to represent the snow, ice and sediment layer temperature profiles.

The resolution of the bulk energy budgets of the mixed layer and thermocline allows to calculate the model prognostic variables, i.e. the mixed-layer temperature and depth, the water bottom temperature, the thermocline shape factor, the ice layer top temperature and ice thickness.

If the snow module is activated, the snow temperature at the atmosphere interface and the snow depth are calculated through an evolution equation that considers the snow precipitation rate within the time step. It should be noted that the model only resolves the energy budget equations and not the water balance. That means that the water volume is kept constant in time. The depth is therefore an input parameter of the model. In addition, two important radiative parameters, surface albedo and light extinction coefficient as well as the surface fetch involved in the calculation of the surface fluxes, need to be prescribed to run the model.

FLake does not model the hypolimnion, the layer under the thermocline which is present or may appear seasonally in stratified lakes and where the water density is the highest. For deep lakes and to get around modeling this layer, the lake depth is set to a maximum value of 50 meters as recommended in the FLake documentation available on the dedicated website (http://www.flake.igb-berlin.de/site/download).

3 ORCHIDEE-FLake coupling developments

3.1 Integrating FLake into ORCHIDEE

The FLake model has been implemented in the ORCHIDEE model, following the same approach proposed by Salgado and Le Moigne (2010), i.e., multi tile approach where separate energy budgets are computed for the vegetation and the lake respective fractions within a grid cell. The developments were done in the SECHIBA routine driving the resolution of the water and energy surface budgets. All the tiles shared the same atmospheric forcing data and exchanges fluxes with the atmosphere according to their respective surface area. The number of lake tiles within a grid cell is variable and can be defined by the user, in order to allow the simulation of different types of water bodies which could present different properties such as depth or radiative properties. Figure 1 presents a general scheme of the ORCHIDEE-FLake coupled model. In the following sections, we present some refinements that were brought in FLake to account for recent improvements in the resolution of the water temperature profile and in the snow model. The parameterization of lakes within the ORCHIDEE grid cell and the choices made for their characterization is then presented.

3.2 FLake model refinements

3.2.1 Shape factor revision

As already noted, in FLake, the temperature profiles in the thermocline layer (but also in the sediment layer not activated in this work) are solved following the concept of self-similarity. The temperature profile in the thermocline is computed through a shape function parametrized by an empirical polynomial equation which satisfies boundary conditions at the top and at the
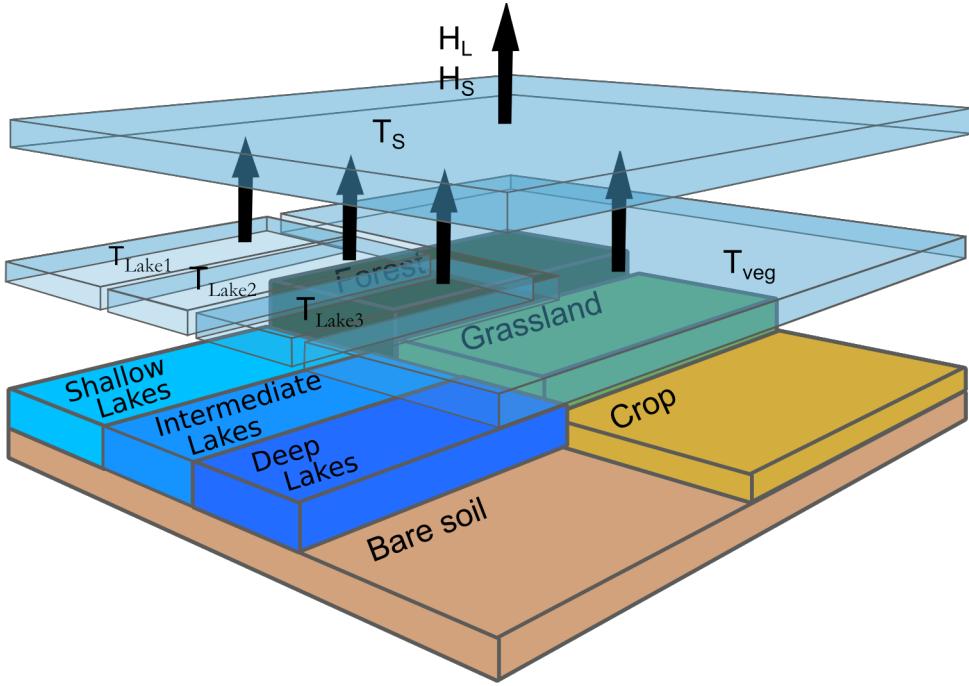
Figure 1. Multi-tile lake representation in the ORCHIDEE land surface model in a 3-tile configuration.

bottom of the thermocline. The equation was proposed by Malkki and Tamsalu (1985) where two cases are differentiated depending if the mixed layer h is developing or retracting following Equation 1.

215
$$\Phi_\Theta(\zeta) = \begin{cases} 1 - (1-\zeta)^3 & \text{if } dh/dt > 0 \\ 1 - 4(1-\zeta)^3 + 3(1-\zeta)^4 & \text{if } dh/dt \leq 0 \end{cases} \quad (1)$$

Φ_Θ represents the normalised temperature depending of the normalised depth ζ as follows:

$$\Phi_\Theta = \frac{T_s - T(z)}{T_s - T_b}$$
$$\zeta = \frac{z}{D} \quad (2)$$

where we define T_s as the lake surface water temperature, T_b as the lake bottom temperature and $T(z)$ is the water temperature at the depth z.

220 Mironov (2008) defined a shape factor C_θ as the integration over the depth of the temperature profile. Equation 3 controls the evolution of the shape factor and allows to make the transition between the two temperature profiles. This factor linearly changes with time and proportionally to a relaxing characteristic time t_{rc} between the two extreme values $C_\theta^{\min}$ and $C_\theta^{\max}$.

$$\frac{dC_\theta}{dt} = \text{sign}(dh/dt)\frac{C_\theta^{\max} - C_\theta^{\min}}{t_{rc}}, \quad C_\theta^{\min} \leq C_\theta \leq C_\theta^{\max} \tag{3}$$

Some issues were reported by Salgado and Le Moigne (2010) who found unrealistic behavior of the temperature profile in
225 spring seasons for long term simulations, with a too large decrease of the bottom temperature. They reveal that the numerical instabilities of the shape factor during winter were playing a role in this issue. On our side, we observed similar instabilities even when no freezing conditions were encountered. Salgado and Le Moigne (2010) proposed to limit the shape factor time evolution to a threshold fixed to $0.01\,h^{-1}$ and to rise up the minimum value $C_\theta^{\min}$ from 0.5 to 0.65. They found that the deep temperature evolution with these changes was more realistic with a large decrease of the numerical instabilities. After some
230 tests, we came to the same conclusions and chose to apply the same corrections to the shape factor parameterization.

3.2.2 Lake snow module

The snow module within FLake which allows to simulate the time evolution of the snowpack when the lake is frozen, was not well tested in its original parameterization (Mironov et al., 2010). The snow temperature is represented following the same approach as the one developed for the ice layer below, i.e., integral heat budget of the ice and snow layers and self-similarity
235 assumption of the temperature profiles. At the snow-ice interface, Equation 4 is used to calculate the snow temperature based on the condition of continuity of the heat flux.

$$T_{snow} = T_{ice} + R_h \cdot (T_{ice} - T_{freeze}) \tag{4}$$

R_h is the ratio of the thermal resistance between the ice layer and the snow layer equal to $\frac{H_S}{H_I}\frac{\Phi'_I(1)}{\Phi'_S(0)}\frac{k_I}{k_S}$ where H_I and H_S are the ice and snow height and k_I and k_S are the thermal conductivity of ice and snow. This formulation can lead to numerical
240 instabilities for high values of R_h. To fix this issue, Pietikäinen et al. (2018) when coupling the German REgional MOdel (REMO, Jacob and Podzun (1997)) with FLake, proposed to set a threshold for the ice layer thickness above which the snow layer can develop. On our side, we did not implement this solution and prefer to decrease the time step of the FLake snow module to prevent the numerical instabilities with the advantage of preserving the water budget. Several global test simulations were performed to adjust the time split factor. As a result, it appeared necessary to go down to time steps of the order of one
245 minute to remove all the instabilities. Therefore, it has been decided to set the time split factor to a value of 50 and to activate it only when the snow thickness is less than 3 cm. Other improvements to the snow module of FLake were proposed by Semmler et al. (2012). The improvements concern the parameterizations of the snow density, the thermal conductivity and the snow/ice albedos. Based on data acquired on Lake Bear in Canada, Semmler et al. (2012) revised the snow density relationship with height which can lead to unrealistic values during the melting period and tested various parameterizations including the use of
250 a constant value of $320\,kg \cdot m^{-3}$. Their results show the benefit of parameterizations accounting for snow aging, but given that the use of a constant value permitted them to better represent the snow and ice cover dynamics compared to the original FLake parameterization, we chose this option as a first step.

The original equation of the snow thermal conductivity $k_{s_{eff}}$ was given by Equation 5 with $k_{s_{max}}$ and $k_{s_{min}}$ the maximum and minimum values of the snow heat conductivity and h_s the snow depth. It was replaced by Equation 6, where $k_{s_{eff}}$ varies between the ice and snow conductivity values according to the snow layer thickness. The bounds of the interval were set to $0.14\,W\,m^{-1}\,K^{-1}$ for the heat conductivity of the fresh snow k_s and $2.29\,W\,m^{-1}\,K^{-1}$ for the heat conductivity of the lake ice k_i and the empirical constant c was fixed to a value of $5\,m^{-1}$, following Semmler et al. (2012).

$$k_{s_{eff}} = min(k_{s_{max}}, k_{s_{min}} + 1.3 \cdot h_s \cdot \frac{\rho_s}{\rho_w}) \tag{5}$$

$$k_{s_{eff}} = k_s + (k_i - k_s) \cdot e^{c \cdot h_s} \tag{6}$$

Following Semmler et al. (2012), Pietikäinen et al. (2018) revised also the calibration of the ice and snow albedos and proposed to change the white/blue ice albedos values from 0.6/0.1 to 0.5/0.3 and the dry/melt snow albedo values from 0.6/0.1 to 0.87/0.77. We implemented the same modifications in our developments.

3.3 Model settings

ORCHIDEE requires the description of the land surface variability in each model grid cell to simulate the energy and biogeochemical processes. As already noted, the standard version uses PFT maps derived from the CCI Land cover product. Since the standard PFT maps do not consider water surfaces, the permanent water bodies from the CCI Land Cover product have been merged with the bare surfaces in the corresponding PFT. The new splitting in SFTs with different energy budgets calculation requires that lake fraction is first removed from the bare soil fraction and transferred to the new lake SFT. For that purpose, new PFT maps excluding lakes for each year of the simulation have been generated. In a second step, we have worked on the definition of the lake fraction within a grid cell and the needs to split it according to the depth, following the recommendations of Bernus et al. (2021). Indeed, one important outcome of our previous work was to quantify the role of FLake parameters in the simulation of the surface variables and to show that depth is especially determinant in the case of shallow lakes. Therefore, we decided to prescribe the number of lake tiles in our model. This new parameter, that has to be set prior to the simulation, allows to define different types of lakes within a grid cell, differing by their input parameters (depth, surface area, radiative properties, . . .). The methodology developed to define those parameters is presented in the following.

3.3.1 Lake surface area and depth

The HydroLAKES database (Messager et al., 2016) was analyzed to parameterize lake surface area and depth in ORCHIDEE. In the study of Bernus et al. (2021), three categories were defined according to depth ranges: shallow lakes with depth below 5 m, deep lakes with depth larger than 25 m and the intermediate class in between (depth range between 5 and 25 m). Figure 2 shows the lake surface fractions (in logarithmic scale), plotted at $0.5°$ resolution, for the three categories of lake depths and aggregated in a single class. The median values of the corresponding depths are shown in the same figure. The
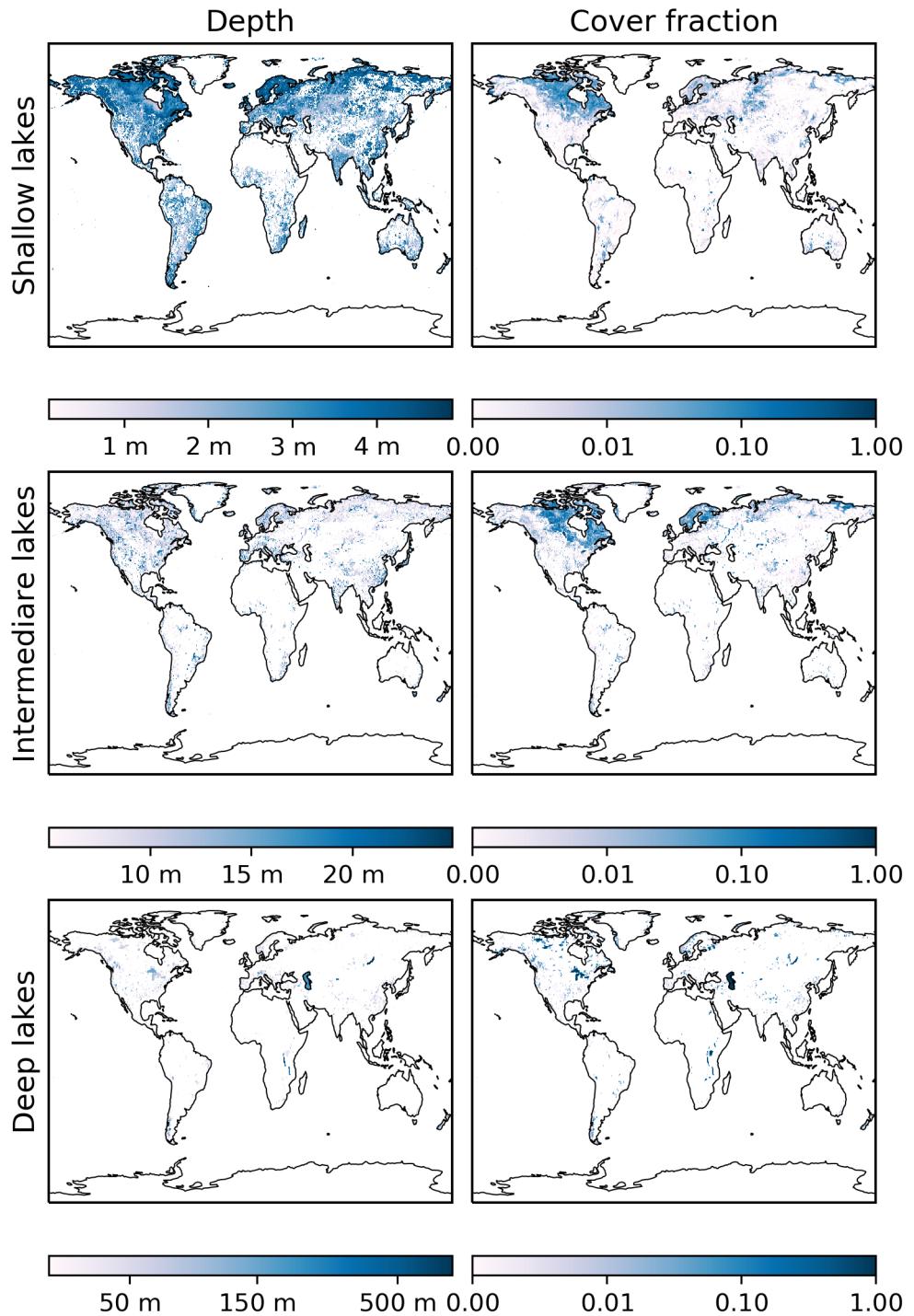
Figure 2. Lake depth (left column) averaged for each lake categories: shallow (below 5 m, intermediate (between 5 and 25 m) and deep (above 25 m), and cover fractions (right column) at a spatial resolution of $0.5°$

higher amount of lake surfaces in the high northern latitudes mainly in northern Canada, Scandinavia and Western Siberia, are clearly highlighted. The lake distribution shows a minimum at the latitudes between $20°$ North and $-30°$ South, in coherence with the arid climates found at these latitudes, produced by the descending branch of the Hadley circulation. The plots of the corresponding depths show a higher concentration of shallow and intermediate lakes in the northern high latitudes, whereas deep lakes look well distributed on all the continents. The major big lakes appear clearly on the plots, as the Great lakes in Canada, Lake Victoria in South Africa, Lake Malawi in South America or Baikal Lake and Caspian Sea in Eurasia. It can be noted that when ORCHIDEE is fully coupled to the IPSL ESM, the Caspian sea as the other seas and oceans is handled by the ocean model NEMO (Madec et al., 2008).

3.3.2 Other lake parameters

FLake requires the prescription of other parameters like the radiative parameters of water (surface albedo and extinction co-efficient) and the wind fetch. For the radiative parameters, the standard values and parameterizations proposed in FLake have been kept. Therefore, an albedo of 0.07 and an extinction coefficient of $1 \ m^{-1}$ was set for water. The fetch parameter which appeared less influent in the sensitivity analysis performed by Bernus et al. (2021), has been defined according to the mean of the fetch values estimated for all the lakes falling in the ORCHIDEE grid cell. More precisely, the fetch for a given lake was approached from its surface extent. Assuming a circular pattern, the fetch was set to the diameter of a circle of equivalent surface, and the effective fetch of the tile was derived from the aggregation of the fetch values of all the lakes falling in the lake tile considered (mean average).

3.4 Performed experiments

ORCHIDEE-FLake was run at global scale forced by the various atmospheric datasets presented in Section 2.1.5 and for two different configurations of lake tiling: the first one when all the lakes are grouped in a single class and the second one, where lakes are grouped in three classes according to lake depth. A control simulation corresponding to the standard ORCHIDEE configuration without lakes (i.e. considered as bare soils) and using only one atmospheric forcing (CRUJRA) was also performed. Table 1 presents the list of all the simulations performed, the atmospheric forcing used and its spatial resolution, the number of lake categories and the way the depth was set for each class (median depth value of all the lakes within the tile, arithmetic, quadratic or harmonic averages). All the simulations were run on the period (1979 - 2016) at a time step of 30 min. The first 20 years (1979 - 1999) were used as spin up to bring the prognostic variables to equilibrium. Therefore, the results were analyzed over the period (2000 – 2016). The name of each simulation is composed of three letters referring to the atmospheric forcing, one number corresponding to the number of lake tiles and between three and five characters referring to the way the depth has been averaged. Then, the comparison of CJR3med to CJR0 allows to assess the contribution of lakes on the ORCHIDEE surface variables. Comparison of CJR3med to CJR1med allows to see the impacts of the number of lake tiles and the way of averaging is seen in the comparison of CJR1med to CJR1mean, CJR1qmean and CJR1hmean. Finally, the comparison of CJR3med to WCR3med, CNC3med, E2O3med and ERA3med, allows to highlight the impact of the atmospheric forcing and

Table 1. List of the ten ORCHIDEE-FLake simulations performed at global scale, differing from the atmospheric forcing used and the number of lake tiles considered in the ORCHIDEE grid cell.

ID	Meteorological Forcing (spatial and temporal resolutions)	Number of lake tiles	Depth calculation method
CJR0	CRUJRA (0.5°, 6 h)	0	No lakes
CJR3med	CRUJRA (0.5°, 6 h)	3	median
WCR3med	WFDEI CRU (0.5°, 3 h)	3	median
CNC3med	CRU-NCEP (0.5°, 6 h)	3	median
E2O3med	E2OFD (0.25°, 3 h)	3	median
ERA3med	ERA5 (31km, 1h)	3	median
CJR1med	CRUJRA (0.5°, 6 h)	1	median
CJR1mean	CRUJRA (0.5°, 6 h)	1	arithmetic mean
CJR1qmean	CRUJRA (0.5°, 6 h)	1	quadratic mean
CJR1hmean	CRUJRA (0.5°, 6 h)	1	harmonic median

the related uncertainties. The performance metrics, i.e. the RMSE, the MSE and its components, used in the evaluation are presented in the Appendix A.

We also compared the aggregation depth of each experience to the depth of the observed lake in HYDROLAKES. To assess the model errors in the comparison to the observations, a preliminary step was to link the lakes from HydroLAKES to the GloboLakes ones. We succeeded to link 964 of the 979 lakes of GloboLakes. The fifteen missing lakes which were not present in HydroLAKES have not been included in the evaluation procedure. When GloboLakes and HydroLAKES were in agreement over a single lake, the depth referenced in HydroLAKES was assigned as the observed depth. But when several water bodies in HydroLAKES corresponded to a single observed lake, we assigned the area weighed mean depth to assess the depth of the observed lake.

4 Results

4.1 Evaluation at global scale

 #### 4.1.1 LSWT

The simulated LSWT obtained in CJR3med configuration have been compared to the observations provided by GloboLakes for the 979 lakes matching ORCHIDEE simulated lake tiles. We recall that CJR3med corresponds to the simulation performed with the meteorological forcing CRUJRA and prescribing three lake tiles differing by their depth (shallow, intermediate and deep lakes). The comparison was done with the lake tile corresponding to the depth category of the observed lake. The RMSE calculated on the 17-year period (2000-2016) are plotted in Figure 3 along with latitudinal and zonal averages. Each dot corresponds to a lake of the GloboLakes dataset. Figure 3 shows that the observations are well distributed globally and sampled a large variety of climates and continents. Their density is well correlated with the lake area distribution plotted in Figure 2. Therefore, the LSWT errors calculated with GloboLakes should be a good representation of the ORCHIDEE-FLake model error. Since GloboLakes does not provide LSWT for frozen lakes, LSWTs below $0°C$ are not used in the RMSE calculations. The results show that most of the RMSEs are ranging between 2 and 4 $°C$. Higher errors can be seen for a few lakes with RMSE values up to 8 $°C$. The latitudinal representation highlights a positive trend of the RMSE, when moving from the temperate to the boreal regions. In the tropics, some large RMSE peak values are observed, which are all the results of individual lakes presenting large errors which impact greatly the zonal average. Because of the low number of lakes at these latitudes, a lake badly simulated has a larger impact on the performance metrics than in the higher latitudes.

4.1.2 Ice phenology

Freezing processes were evaluated against the data provided by the Global Lake and River Ice Phenology database over the lakes observed in the northern latitudes. Since the observations provide the dates of the full ice end, start and duration of a complete ice coverage, we compared with the same simulated dates and did not consider ice free periods (if any) within this interval. In Table 2, the mean, quartiles and median of the differences between simulated and observed three stages of the ice phenology are given in days unit for the CJR3med simulation. The results show that the ice start errors are all negative and the ice end errors are all positive. That means that the lakes are generally freezing earlier and melting later in the model compared to the observations. Therefore, the simulated ice duration is larger. The mean bias on ice start is about -11 days (median value), +21 days for ice end, leading to an ice duration overestimated by 32 days in average. The RMSE values are around 24 days for both ice start and ice end and reach 40 days for the ice duration error. The interquartile range is larger for ice start (19 days) compared to ice end (10 days) highlighting the fact that the ice start errors are more scattered. This also explains why the differences between the median and mean values are larger for ice start compared to ice end.
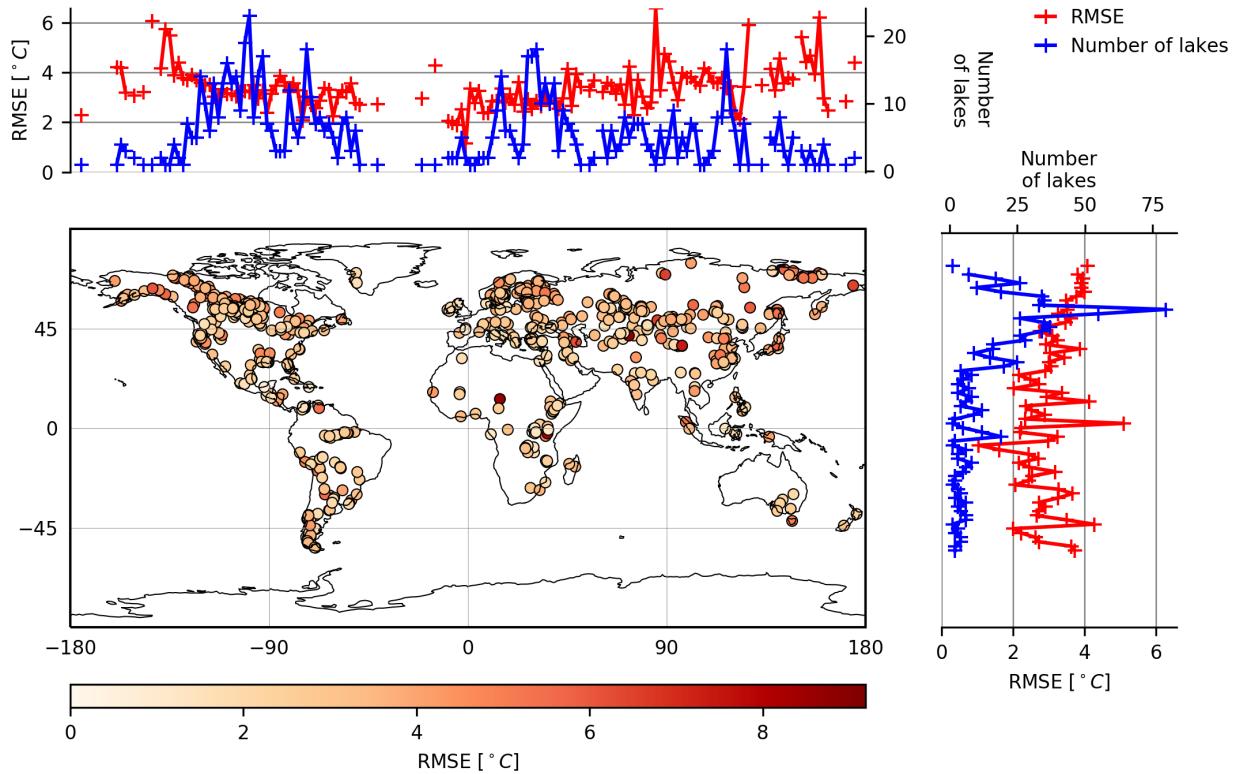
Figure 3. LSWT errors (RMSEs) simulated by ORCHIDEE-FLake in the CJR3med configuration run at $0.5°$ resolution. Longitudinal (top) and latitudinal (right) averages of the RMSEs (in red) and number of lakes observed in GloboLakes (in blue) are plotted along the map.

Table 2. Statistical metrics (simulations - observations) of the lake ice phenology (ice start, end and duration) obtained in the CJR3med configuration

Statistical metrics	Ice start (days)	Ice end (days)	Ice duration (days)
Median	-11.46	+21.5	+32.35
Q1	-4.9	+26.7	+43.5
Q3	-23.9	+16.3	+24.1
Mean	-16.46	+22.1	+36.0
RMSE	24.1	24.0	40.5

4.2 Sensitivity to meteorological forcing

4.2.1 Analysis on specific lakes

Model evaluation has been performed at the local scale for various lakes sampling a large variety of climates and lake properties,
in particular their mean depth. Figure 4 presents the comparison of simulated LSWTs to observations for a selection of eight
lakes. Among those, Victoria and Titicaca are located in tropical regions, Erie, Superior and Baikal in boreal ones, Constance,
Tahoe and Tai in more temperate ones. The lake Tai is a shallow lake, the others being deeper (mean depth of each lake is noted
on the respective plots). Another criteria for the choice of these lakes was their spatial extent, covering at least one ORCHIDEE
grid cell at 0.5° resolution, enabling a better comparison with the observations. The LSWT time series (simulated and observed
over the year 2010 chosen as an example) are plotted in Figure 4 along with the MSE errors decomposed into its 3 components
(Kobayashi and Salam, 2000), i.e., squared bias (SB), squared difference between standard deviations (SDSD) and lack of
correlation weighted by the standard deviation (LCS), all represented in bar plots. We recall that the metrics are calculated only
for positive values of LSWT, since GloboLakes observations do not consider frozen conditions.

Figure 4 shows that the seasonal variations of LSWT are generally well represented for all the lakes. The MSE values are
ranging between 0.9 and 27.7 $°C^2$ (i.e., RMSE between 0.95 and 5.3 $°C$) but can show large variations according to the
atmospheric forcing. The MSE decomposition show that the model errors are either explained by biases or by deficiencies
to simulate the fluctuation patterns, and less by failure to model the amplitude of the seasonal cycle, since SDSD is always
much lower than the SB and LCS components. Among all the lakes, Lakes Tai and Tahoe present the lowest errors and Lake
Titicaca the largest ones for at least three of the forcing used. It is interesting to note that the results do not vary much among
the forcings for a same lake, except for the two tropical lakes, Victoria and Titicaca, the latest showing the largest variability.
Indeed, RMSE of Lake Titicaca is equal to 0.9 $°C$ when CRUJRA forcing is used and can be up to 5 $°C$ for CRU-NCEP or
WFDEI-CRU forcings. It is the same for Lake Victoria where RMSE is around 0.95 $°C$ in the E2OFD simulation and up to
3.5 $°C$ in the WFDEI-CRU one, all these errors explained mainly by biases.

For the other lakes, the errors are more homogeneous among the forcings, even if significant differences appear. If we
consider the three temperate lakes, we can see that the annual surface temperature amplitude is larger than for the tropical ones,
between 15 and 20 $°C$. Lake Tai is well represented with MSEs below 8 $°C^2$ (i.e, RMSE below 2.8 $°C$) for all the forcings.
For the two other temperate lakes (Tahoe and Constance), the model fails to simulate the annual amplitude which is slightly
underestimated, and the cooling at the end of summer is delayed whatever the forcing used. For the Lake Constance, the SDSD
component can reach 4.5 $°C^2$. Such errors could result from parameterization errors, such as depth errors which are known to
influence greatly the seasonal LSWT pattern.

The results obtained for the boreal lakes do not show the same features: the cooling after summer is generally well reproduced
whatever the forcing used but the warming in spring looks too rapid for Lake Erie and Lake Superior and too late for Lake
Baikal. Frozen processes show also large deficiencies in all simulations, with Lakes Superior and Erie freezing in winter,
contrary to the observations and an overestimation of the frozen season for Lake Baikal which is freezing by about one month
in advance with CRU-NCEP forcing and melting one month too late, preventing to simulate correctly the temperature time

evolution in spring. In these two cases, the simulations performed with E2OFD forcing show the best results in terms of MSE. The fact that the duration of the freezing period is largely overestimated whatever the atmospheric forcing could be the result of parameterisation errors of the cryospheric processes that will be investigated in the following.

4.2.2 Global scale analysis

390 Figure 5 presents the performance metrics calculated on all the lakes observed in GloboLakes (i.e., 979 lakes all around the world). The comparison is done at the lake level with the corresponding ORCHIDEE lake tile the lake falls in (given its mean depth). The statistics were calculated for the seventeen years of the whole simulation period (2000-2016). The metrics shown here are the RMSE, the unbiased RMSE, and the three normalized components of the MSE, i.e., SB (with the sign), SDSD and LCS, in terms of fractions, calculated for each simulation forced by the five atmospheric forcings used. We recall that the sum

395 of the 3 components is equal to unity.

The results show that the simulations forced by the meteorological reanalysis with the lower resolution (i.e., $0.5°$) lead to the worst performances (higher RMSEs between $3.18\,°C$ for CRUJRA, $3.16\,°C$ for CRUNCEP and $2.82\,°C$ for WFDEI-CRU). When higher resolution forcings are used (E2OFD and ERA5), the RMSEs decrease to $2.81\,°C$ and $2.69\,°C$. An important part of the errors is therefore explained by the resolution and accuracy of the meteorological forcings. It should also be noted

400 that ERA5 is based on the ECMWF atmospheric model which includes the HTESSEL land surface model accounting for lakes, unlike the other atmospheric products. This feature could lead to atmospheric surface variables of better quality over lakes and explain that the simulated lake surface temperatures are closer to the observations.

However, the decomposition of the MSE is quite similar among the forcings on average. For half of the lakes, LCS component represents at least 45% of the MSE error for CRUJRA and ERA5 simulations and up to 65% for the E2OFD simulation.

405 The SDSD component is very minor since it is less than 20% of the MSE for more than 75% of the observed lakes. Thus, the LSWT errors are firstly explained by the model biases and deficiencies to reproduce the seasonal patterns, confirming what was already noted for the individual lakes analyzed previously. In Figure 5, we can see that the biases are positive for all the forcings, meaning that the model generally overestimates the surface temperature. The signed SB (median values) obtained with E2O 3med is the closest to 0. Compared to the total RMSE that can reach up to $10\,°C$, the unbiased RMSEs are reduced

410 and do not exceed $6\,°C$.

The sensitivity of the simulated ice phenology to the meteorological forcing has been evaluated and the results are plotted in Figure 6. The whisker plots show that the impact of the meteorological forcing is lower on the prediction of the start of lake freezing (ice start) than on the prediction of the end of the freezing period (ice end), and consequently on the ice duration where the two errors cumulate. It can be seen also that the errors are smaller for the two simulations performed with the higher

415 resolution forcings (E2OFD and ERA5) with a significant advantage of E2OFD especially for the simulation of ice end. The ice start, ice end and duration median errors are respectively equal to -6, 11 and 20 days with E2OFD whereas they are up to -14, 25 and 41 days with the CRUNCEP forcing.
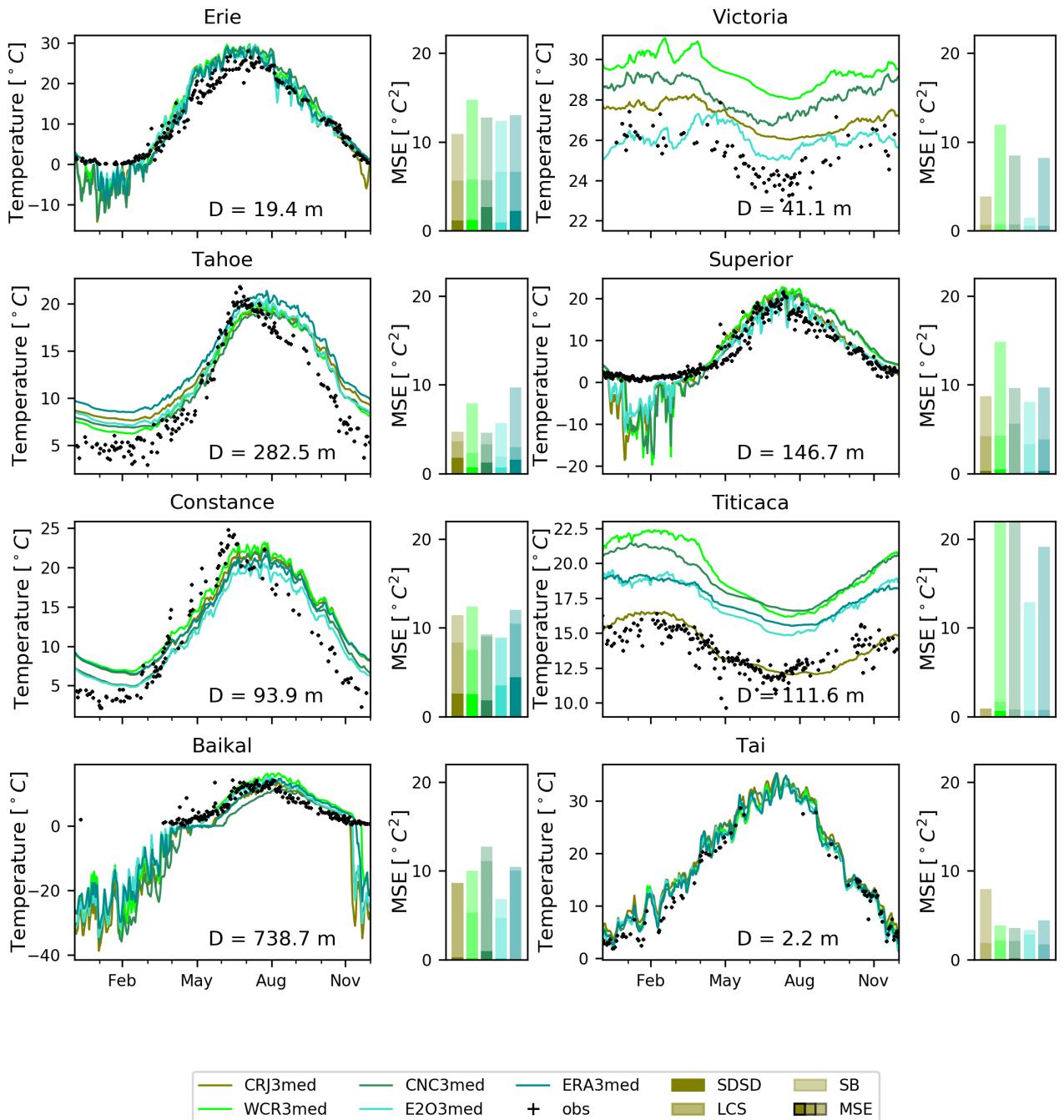
Figure 4. Example of time series of the daily mean lake surface temperature simulated for eight lakes with five different meteorological forcings compared to GloboLakes observations for the year 2010. The barplots represent the decomposition of the MSEs in terms of bias (SB), lack of correlation (LCS) and difference of standard deviations (SDSD) (represented with different shadings) between the model and the observations.
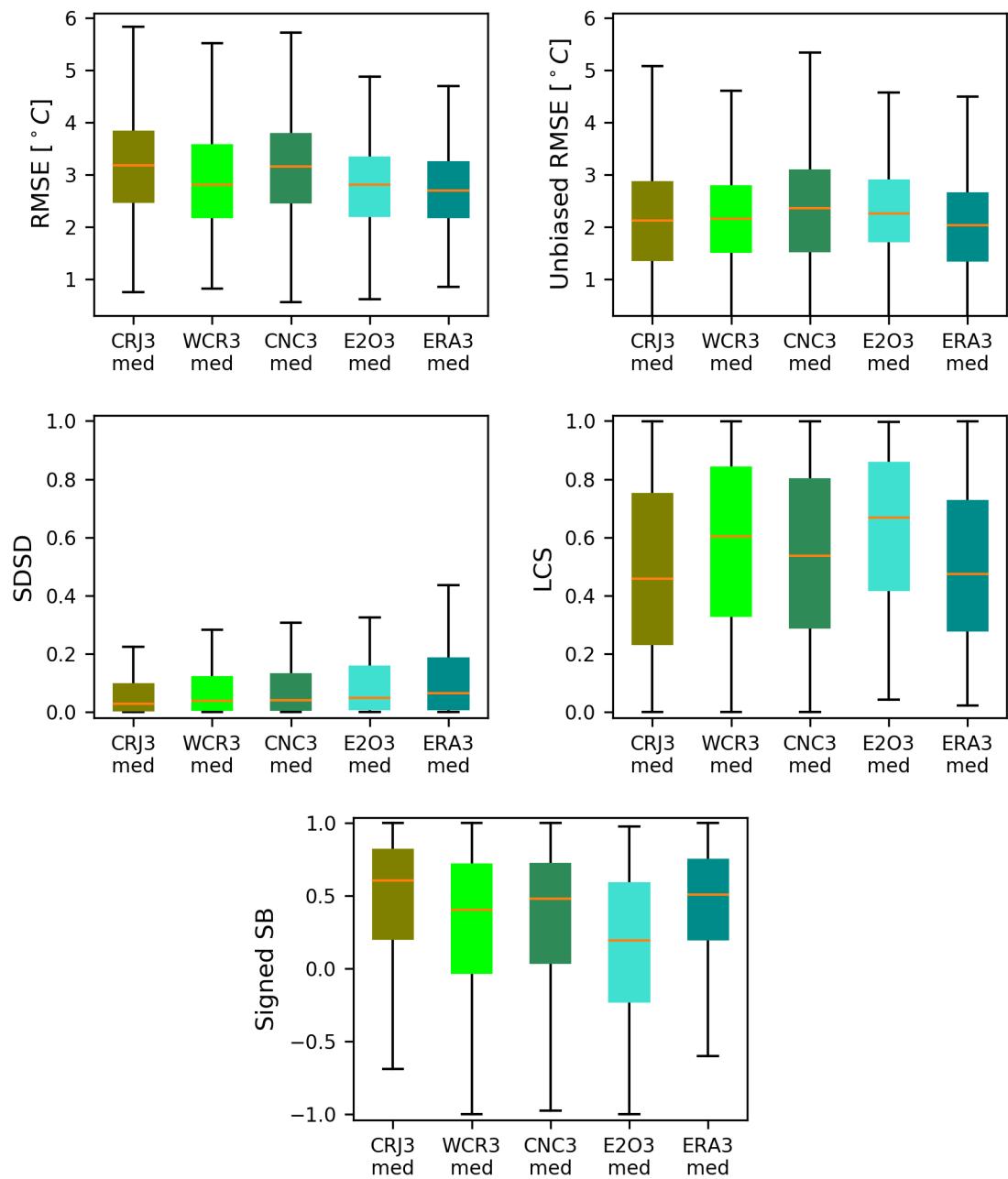
Figure 5. LSWT errors obtained on the 979 lakes observed in GloboLakes for the five simulations CRJRA3med, WCR3med, CNC3med, E2O3med and ERA3med. RMSE, unbiased RMSE, LCS, SDSD and sign SB are represented in boxplots. Median values are represented by orange lines, and quartiles by the edges of the color box.
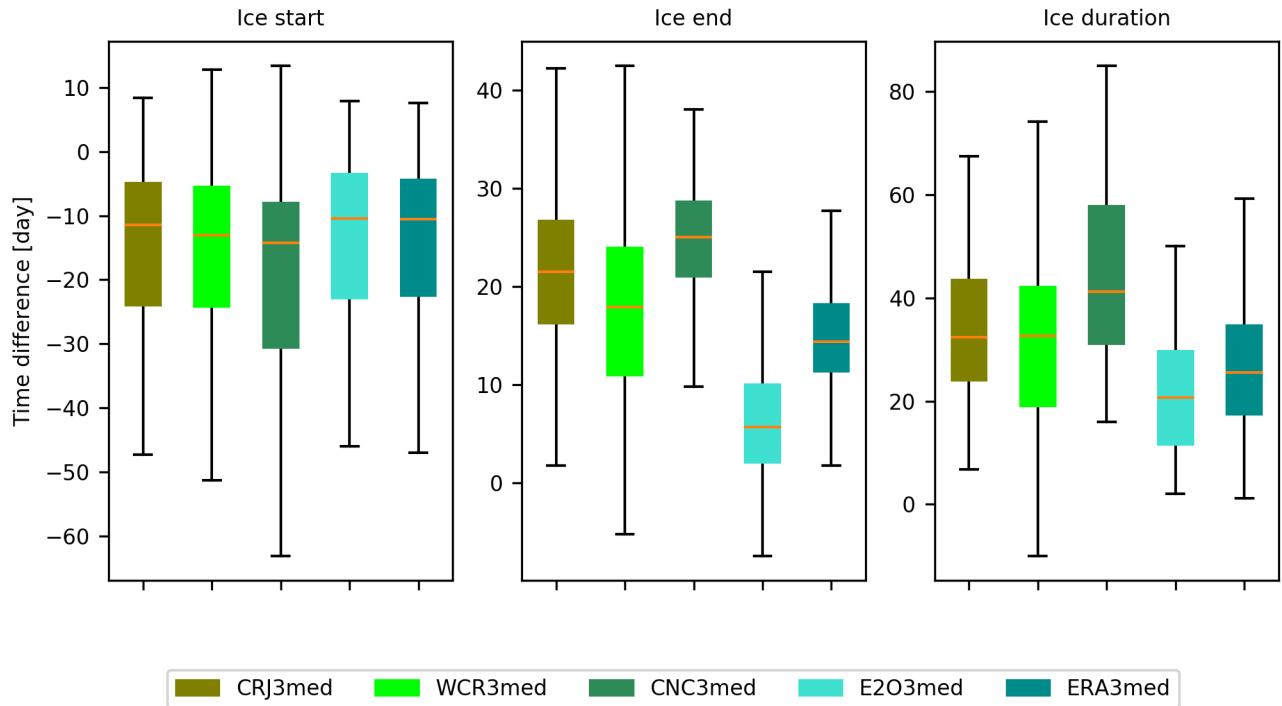
Figure 6. Boxplot of the model errors on the start, end and duration of the ice period for the five simulations CJR3med, WCR3med, CNC3med, E2O3med and ERA3med. Median values are represented by orange lines, and quartiles by the edges of the color box.

4.3 Sensitivity to lake depth parameterization

To understand the impact of the depth parametrization, we analyse the differences between the real depth of the 979 observed lakes and the depth prescribed in the corresponding ORCHIDEE tile at a grid resolution of $0.5°$.

Figure 7 presents the distribution of the errors between the depth of the observed lake and the depth prescribed in the corresponding ORCHIDEE tile, for the five model configurations, i.e., median depth value calculated in a 3-lake tile configuration (3med) and in the 1-lake tile configurations (1med, 1mean, 1hmean and 1qmean). The histogram of the errors and the cumulative distribution functions (CDF) are both plotted. Table 3 gives the median and the ninth decile of the depth error distribution.

The histogram of the errors are well centered around 0 for the 3-tile configuration and the 1-tile one when the median average is performed. The 3 other ways to aggregate the lake depths (arithmetic, quadratic and harmonic means) show negative biases (the median ranging between -0.1 m and -0.7 m) illustrating the fact that grouping all the lakes of a same mesh in a single tile, lead to a slight underestimation of the representative depth. This is probably the result of the worldwide lake distribution previously presented in Figure 2, showing that shallow lakes are more numerous and cover a larger extent than deeper lakes.
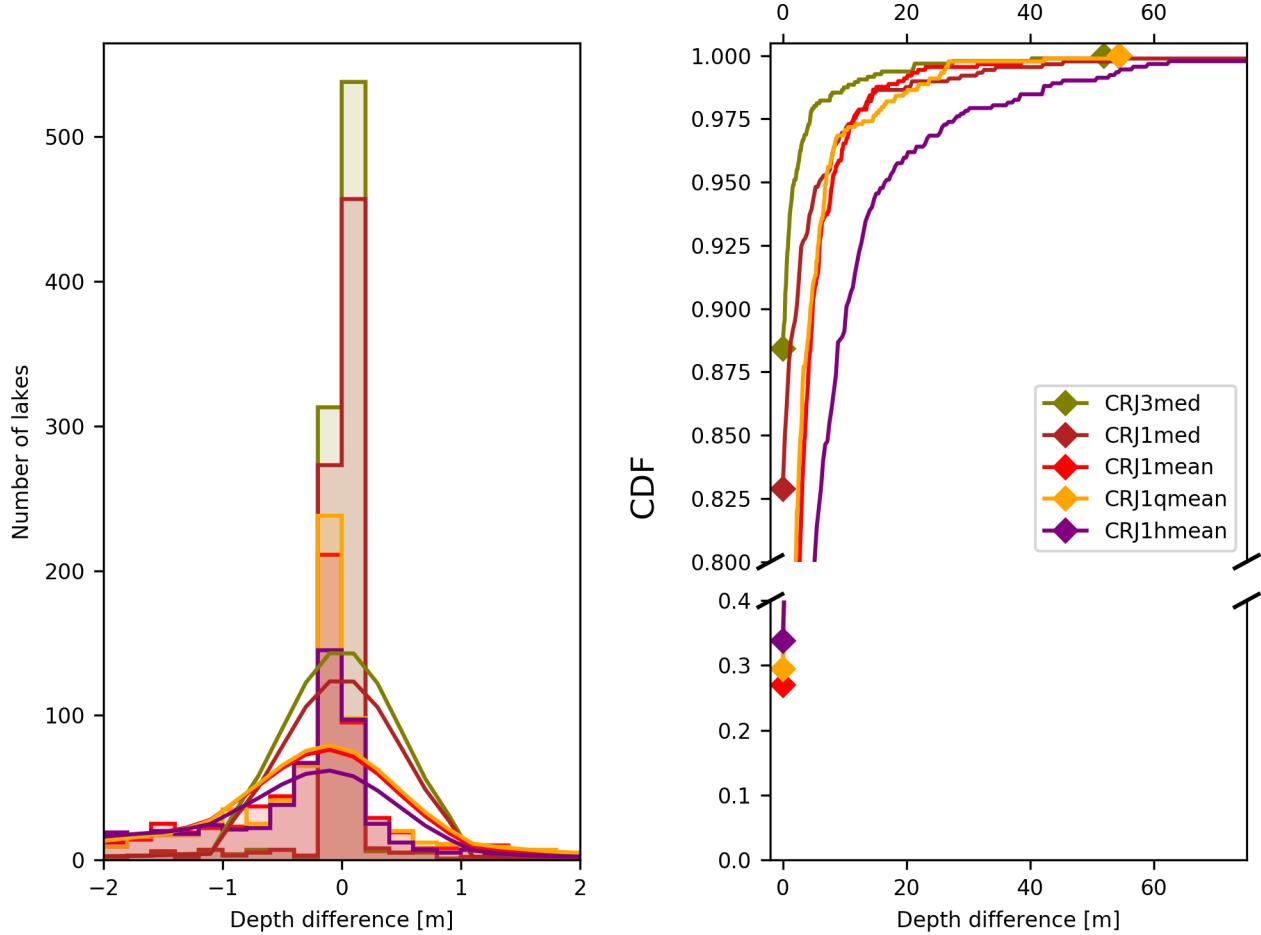
Figure 7. The histogram (left panel) and cumulative density function (right panel) of the depth errors between the observed and simulated lake for each depth parameterization strategy. In the left panel, the lines correspond to the smoothing of the histogram with a Gaussian window.

Table 3. Statistical metrics of the depth error distributions according to the depth parameterization strategy used: the median and 9th decile values are given in meters.

Experiment	Median (m)	9th decile (m)
CRJ3med	0.	0.3
CRJ1med	0.	2.2
CRJ1mean	-0.17	4.9
CRJ1qmean	-0.09	4.4
CRJ1hmean	-0.73	10.2

This is confirmed by the CDF curves plotted in Figure 7 showing that the configuration with 3 tiles presents the lowest errors on the depth parameterization, and the median appears to be the best way of depth aggregation in the 1-tile configurations. Indeed, we have seen that about 90% of the observed lakes have an absolute error less than 0.3 m in the 3-tile configuration and less than 2.2 m when the depth parameter is set to the median value within the lake tile. In comparison, 90% of the observed
435 lakes have an absolute error less than 4.4 m with the quadratic mean and less than 4.9 m with mean depth estimation approach. The harmonic mean is the worst approach since 10% of the observed lakes have an absolute depth error larger than 10.2 m.

In Figure 8, the LSWT errors obtained for all the simulations using the 1-tile configuration are compared to the 3-tile simulation already presented in Section 3.4 . The total and unbiased RMSEs, SDSD, LCS and the signed bias are plotted as box plots. Compared to the previous sensitivity analysis to the atmospheric forcing, the errors are slightly lower as well as the
440 dispersion among the simulations. The RMSE values are generally lower than 6 $^{\circ}C$ with a median value between 3.1 $^{\circ}C$ and 3.2 $^{\circ}C$. The components present similar order of magnitude among the simulations with positive bias and lack of correlation (LCS) explaining most of the MSE, as already seen in the 3-tile configuration. This result was expected since the impact of the lake depth parameter in FLake, impacts mostly the pattern of the seasonal cycle of LSWT and to a lesser extent, its amplitude.

The ice phenology sensitivity to the number of lake tiles and depth parameterization strategy is presented in Figure 9.
445 Here also, the results show less impact compared to the sensitivity obtained with the meteorological forcing. In contrast, it appears that the ice start errors are reduced by half in the 1-tile configurations compared to the 3-tile one. In the four 1-tile simulations, ice start errors are around 5 days whereas it is equal to 10 days in the CJR3med simulation. However, the variance is quite similar, the 3-tile simulation shows a clear bias not observed in the 1-tile simulations. Ice end and duration errors are similar among the five simulations, equal to about 22 days and 27 days respectively, comparable to what is obtained in the
450 CJR3med simulation. The improvement observed in the simulation of ice start in the 1-tile simulations compared to the 3-tile one, despite the increased depth errors, can be explained by the model biases. As a matter of fact, we highlighted in Section 4.2.1 the systematic underestimation of ice start time: LSWT cooling at the end of summer is too rapid and the lake freezing is earlier compared to the observations. If we consider that the shallow lakes represent a larger fraction of the lakes which are most likely to freeze and that the depth parameter of this class is the most impacted in the 1-tile configuration, it is clear
455 that parameter errors are compensating model deficiencies. The depth of shallow lakes, is most likely to be overestimated in the 1-tile configuration. The deeper the lake, the later the start of freezing. Therefore, the depth overestimation will help to
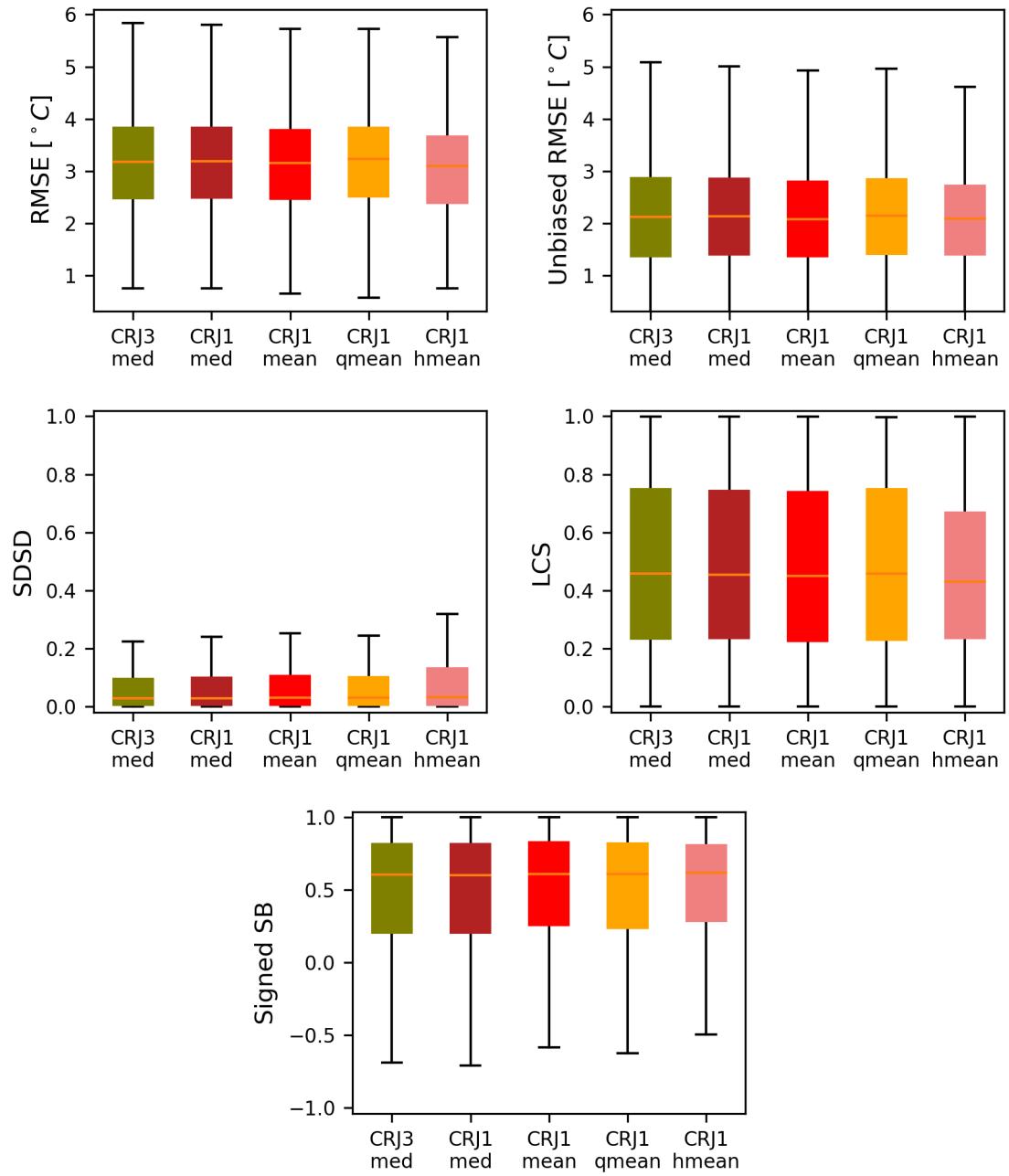
Figure 8. LSWT errors obtained on the 979 lakes observed in GloboLakes for the five simulations CJR3med, CJR1med, CJR1mean, CJR1qmean and CJR1hmean. RMSE, unbiased RMSE, LCS, SDSD and SB are represented as scatterplots (left panel) in comparison to the ensemble mean and in boxplots (right panel)Median values are represented by orange lines, and quartiles by the edges of the color box.

compensate the too early cooling. The impact of this error is less visible on the ice end because the depth parameter is less influential during the ice melting process (Bernus et al., 2021).
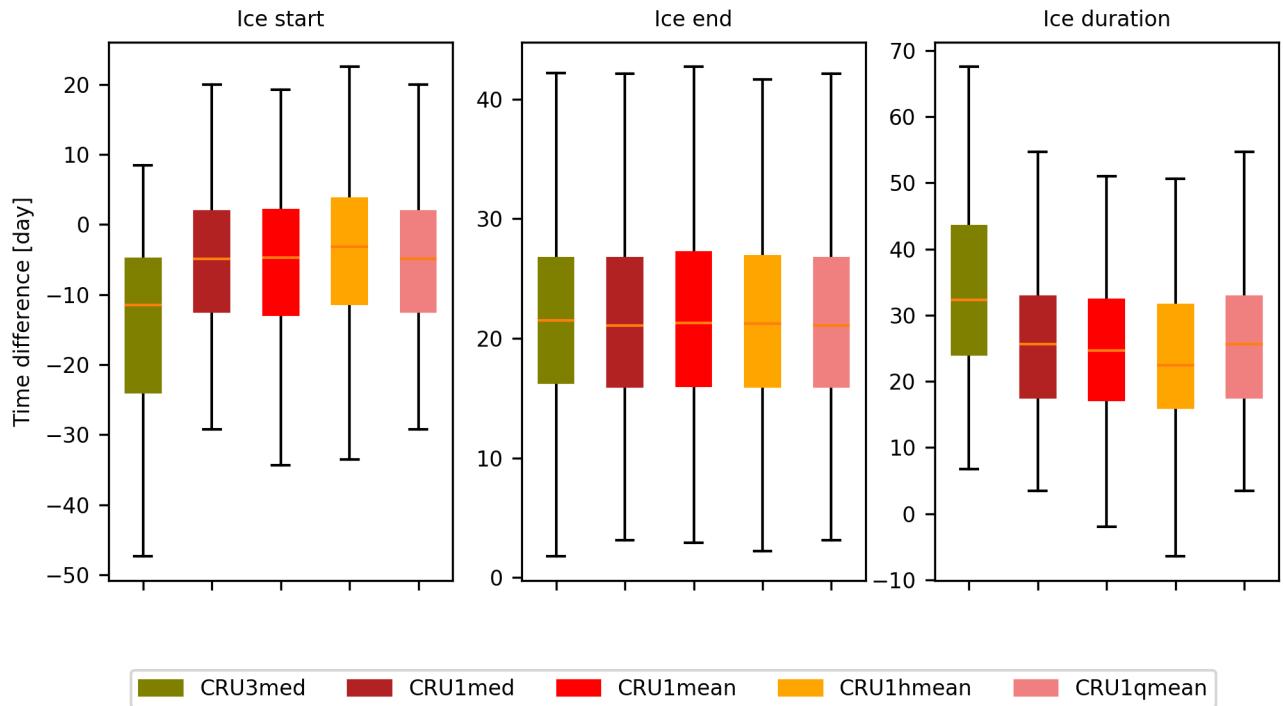
Figure 9. Boxplot of the RMSE errors on the start, end and duration of the ice period for the five simulations CJR3med, CJR1med, CJR1mean, CJR1qmean and CJR1hmean. Median values are represented by orange lines, and quartiles by the edges of the color box.

4.4 Energy balance impacts of lakes in ORCHIDEE

460 To analyse the impact of representing lakes on the global energy and water budgets, the two simulations CJR0 et CJR3med were compared at the global scale. The differences (CJR3med-CJR0) in terms of surface temperature, latent and heat fluxes averaged over the grid cell, were calculated for the four seasons of the year and plotted in Figure 10. Winter period is defined as the average of the December, January and February months, summer as the average of June, July and August.

The comparison of the two simulations (with and without lakes) shows interesting features. First, the surface temperatures
465 appear warmer during a large part of the year, except for summer where the regions presenting large fractions of lakes of intermediate depth, appear colder. In the high latitudes during the winter and autumn seasons, the temperature is generally warmer for all kind of lakes. For the deeper lakes, this difference can reach 10 $°C$. It is the case for Lake Baikal whereas for the Michigan and Superior lakes, the difference is between 5 and 7 $°C$. In the regions covered by smaller and shallower lakes in Scandinavia and Canada, this difference remains below 2 $°C$ due to the lower lake fraction and smaller smoothing effect
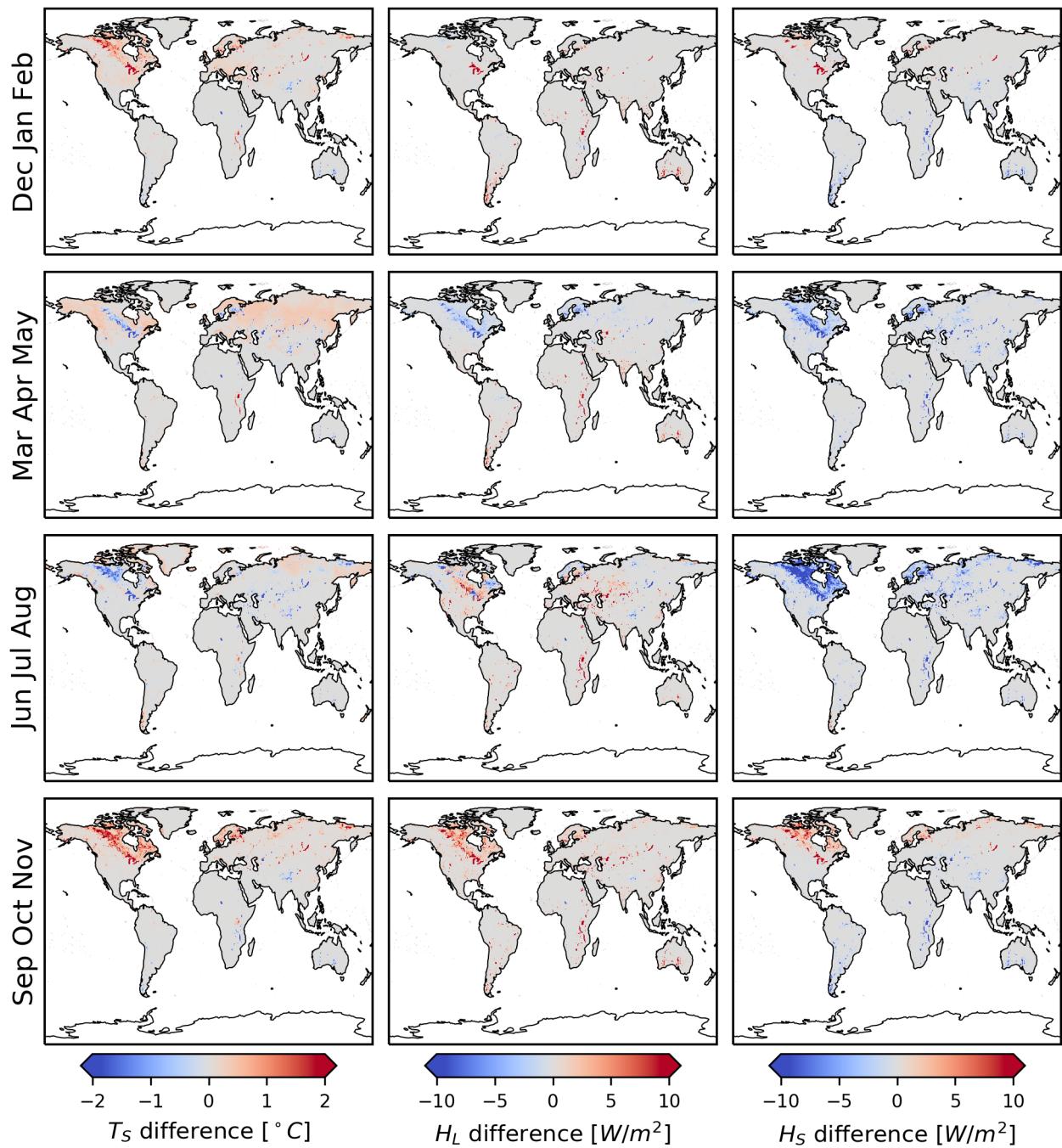
Figure 10. Differences (CJR3med - CJR0) in the surface temperature (first line) latent heat flux (second line) and sensible heat flux, simulated with and without lakes, averaged over 3-month periods starting from December (one per column).

470 of shallower lakes. In spring, the deep lakes as the great American lakes or Lake Baikal are cooler than their environment and explain the differences of temperature between the two simulations by 5-6 $°C$ and 3-4 $°C$ respectively. However, the regions where the lakes are shallower remain warmer in the lake simulation because of their lower thermal inertia (but only by 0.5 $°C$). In summer, most of the lakes are cooler in the northern latitudes compared to a neighbouring bare soil and the differences between the two simulations are always positive. In tropical latitudes, the pattern is different. During the humid season between

475 October and April, the simulation with lakes is warmer at the end of the humid season by 1 or 2 $°C$ (in Africa during the spring season) and cooler during the end of the dry season by 0 or 3 $°C$ (in Africa).

Averaged at yearly time scale, the simulated global evaporation is 1.07 mm/d with lakes and 0.82 mm/d without lakes. Focusing on the summer period, the global evaporation rises up to 1.72 mm/d with lakes and to 1.50 mm/d without lakes when lakes occupy at least 20% of the cell fraction. Then, the contribution of lakes rises the summer evaporation by 0.22 mm/d. For a

480 grid cell completely covered by water, i.e., lake fraction equal to 1, the evaporation difference can reach 0.31 mm/d. In tropical areas, during the summer period and for latitudes below $30°N$, this difference is larger and equal to 1.16 mm/d when the lake fraction is above 20% and amount to 1.6 mm/d when the lake fraction is equal to 1.

In terms of fluxes, the latent heat flux is generally stronger everywhere during part of spring and summer as it is showed by the rise of the mean evaporation rates. In spring and in the northern latitudes, the latent heat flux is lower in the simulation with

485 lakes, because of the freezing processes which prevent the lake evaporation. Compared to a bare soil that start to evaporate earlier in ORCHIDEE, the lake evaporation is lower in early spring. For example, the latent heat flux of Lake Baikal is between 25 and $50\,W/m^2$ less than the bare soil evaporation and it is about the same for Lake Superior (between 20 and $40\,W/m^2$). The freezing duration is longer for a lake than for a bare soil. The sensible heat fluxes are generally weaker everywhere during part of spring and summer because of the larger latent heat fluxes in the simulation with lakes.

490 ## 5   Discussion and conclusion

This work presents how lakes energy balance is now represented in the ORCHIDEE land surface model, thanks to the FLake lake model. The methodology and the model evaluation at global scale were presented. Several simulations were performed to evaluate the model sensitivity to the meteorological forcing and to the depth parameterization of the lake tile. The evaluation of LSWT and ice phenology against the GloboLakes and Global Lake and River Ice Phenology databases, showed that the

495 lake surface temperature is quite well simulated in this new ORCHIDEE version. Compared to the observations, LSWT errors range between 2 and 4 $°C$, even if in some few cases, the errors can reach 8 $°C$. These results are comparable to what was obtained by previous works, such as Dutra et al. (2010) who obtained RMSEs around 3.75 $°C$ in the high latitudes ($> 40°$) and around 2.25 $°C$ for low latitudes ($< 40°$), as median values. The lake ice phenology evaluation showed that the ice/snow coupled processes are also fairly well simulated with ice duration errors explained partly by start of freezing errors. Here also,

500 compared to previous studies, the results obtained are quite similar. As an example, Dutra et al. (2010) obtained an ice start error of 30 days (in advance) for 75% of lakes after optimisation of the depth parameter against LSWT observations. The expected impacts of lakes on the near-surface atmospheric variables have been assessed by comparing the grid cell averaged

surface temperatures and fluxes simulated by ORCHIDEE, considering lakes energy balance or not (in that case, lakes were considered as bare soils). The results show larger evaporation rates at the annual scale (around 0.25 mm/d) and smoother LSWT seasonal amplitudes. The impacts of these added fluxes on the atmosphere should be significant at regional scales, in particular on precipitation predicted fields. As an example, Van de Walle et al. (2020) showed that Victoria lake evaporation explains up to 50% local precipitation and storm events. The addition of lakes in the IPSL climate model is expected to produce similar impacts, helping to better reproduce the near-surface variables, the precipitation regimes and the extreme events.

5.1 Atmospheric forcing uncertainties

A large part of the LSWT errors is explained by the atmospheric forcing uncertainties. We have examined in more detail the lakes which present the largest RMSEs. Among the 969 lakes studied, only 15 present large RMSEs, larger than 6 $°C$. For those lakes, we observed strong biases in the atmospheric temperatures and downwelling longwave radiation (whatever the reanalysis product) which explain the largest part of the error. If we consider the whole dataset, our results show the large sensitivity of LSWT to the atmospheric conditions and the large uncertainties of the reanalysis products. When looking to the RMSE decomposition, biases were explaining more than 50% of the error for more than 50% of lakes. It was particularly the case for the tropical lakes where the LSWT seasonal amplitude is lower than for higher latitudes conditions. The model sensitivity to the atmospheric conditions is also very strong during the intermediate seasons for the lakes which are subject to freezing. Because of the non-linearity of these processes, a small bias on the air temperature can completely change the temperature evolution of the lake, and greatly influence the seasonal pattern and the ice phenology. The large sensitivity of the ice start and ice end times to the atmospheric forcing shown in our simulations, is a good demonstration of this effect. Therefore, it appears very necessary to use various atmospheric forcing products in the model evaluation, since they all contain biases, which on top of that, present regional features. These biases are the result of the different atmospheric models and datasets used in the generation process, which consists in merging model simulations and multiscale observations through data assimilation techniques. It should be noted that the land surface models used in this process do not account for lakes except for the ERA5 forcing and are constrained by static land cover maps which are not updated regularly. Biases are also linked to the spatial and time resolution of the datasets. It has been shown that LSWT errors were smaller when the two high resolution forcings are used (E2OFD and ERA5). Increasing the spatial resolution improved LSWT RMSEs by 0.5 $°C$. Grid resolution is indeed an important parameter for the modelling of land surfaces which have very different functioning than the rest of the environment. Moreover, the atmospheric conditions above a lake surface are different than in the surroundings, complicating greatly the model evaluation at global scale. The use of various reanalysis products is, therefore, the only way to separate forcing and model errors.

5.2 Model parameterization

Systematic model errors appeared among the simulations and could be the result of parameterization errors. Sensitivity to depth parameterization was highlighted in previous works. Our simulations did not show a large impact of the depth parameterization at global scale on the LSWT error statistics, but the impacts on the ice phenology was shown to be significant, especially in

the determination of the time of the start of freezing (ice start). As already said, the influence of lake depth on lake freezing is stronger than on ice melting which is more driven by the atmospheric conditions and the amount of snow than directly on the lake features. These results are in line with those obtained by Choulga et al. (2019), who previously showed the impact of the lake depth on the ice phenology.

540   Scaling issues may explain also part of the errors: FLake is a 1-D model and therefore, do not consider the spatial variability of the within lake thermal processes and properties, nor the temporal variability of the parameters. In particular, the lake depth is a constant parameter in FLake whereas it is clear that in reality, it varies spatially and temporally, and so do LSWT and ice phenology. Lakes, especially those subject to anthropogenic management as reservoirs, see their surface and water level changing throughout the year. The averaged depth and surface area referenced in the databases do not report these changes.

545 For example, Lake Chad in HydroLAKES has a surface area doubled compared to the current value and the average depth is probably erroneous as well. Besides, if we consider the ice observations, the time of ice start refers to a complete ice coverage whereas it can take some days before the lake is completely frozen. In contrast in FLake, because of the 1-D assumption, the lake surface is considered completely frozen when the surface temperature falls below $0°C$. Then, the ice layer starts to grow and so abruptly will vary the related radiative parameters. For big lakes with large surface areas, this scaling error may

550 be important. Similar conclusions were drawn by Pietikäinen et al. (2018) who identified two factors explaining the delay in their simulated ice end and the advance of ice start with the coupled FLake-REMO model. Because lakes are freezing too early, the snow accumulation can start earlier. An overestimation of the snow cover may explain the delay in the timing of the end of freezing and the overestimation of the ice duration period. To overcome this issue, the parameterization of the freezing processes could be revised by accounting for a partial ice cover as proposed by Choulga et al. (2019) or Garnaud et al. (2022),

555 for example. In these two studies, the ice coverage is defined according to the ice layer thickness and so do the radiative properties. Such a modeling could help to slow down the freezing processes in FLake and improve the modeling of the surface temperature and fluxes. The prescription of a constant average value of the snow conductivity in our simulations could also play a role in the deficient representation of the dynamics of the ice/snow melting. Moreover, processes as the snow erosion and transport by the wind are not modelled and may lead to an overestimation of the snow accumulation in the simulations.

560   Scaling and parameterization issues may affect other parameters like the radiative ones: albedo and extinction coefficient which were both prescribed as constant values in space and time. This assumption may also impact the model performances. Recent databases can be used to solve this issue, based on satellite measurements. Albedo products could be used to spatialize the albedo parameters. The extinction coefficients are also known to be key (Heiskanen et al., 2015). For example, Zolfaghari et al. (2017) demonstrated the improvements brought by constraining the extinction coefficient of Lake Superior using FLake,

565 based on remote sensing observations.

    Missing processes could also explain part of the errors such as the impact of water salinity, which could present seasonal and interannual variations linked to human activities and water level changes. In our evaluation, among the fifteen lakes presenting the largest errors, three of them are saline lakes (Issyk-Kul, Urmia and Natron) with high salinity levels, and which can present salt crust regularly (Zavialov et al. (2018), Sima et al. (2013), Tebbs et al. (2013)). Salinity impacts surface albedo

570 and evaporation (which is lower compared to those of freshwater), and consequently the surface temperature and its seasonal

evolution. As reported by Ahmadzadeh Kokya et al. (2011), the temperature rising during the summer can be more rapid for saline lakes, as they observed for the Urmia lake in Azerbaijan. The consideration of salinity could therefore improve the modeling of the energy balance in the regions subject to high evaporation and also the modeling of the cryospheric processes in the cold ones.

5.3 Future work

Our results have shown the potential of ORCHIDEE-FLake coupled model to represent the energy balance of lakes, the surface temperature and fluxes at the atmosphere interface. The model performances are satisfactory, and the various sources of errors are well identified. The flexibility of our developments allowing to prescribe a variable number of lake tiles within a model grid cell, open new perspectives for improvement. First, we intend to study the spatial and temporal variability of the lake radiative parameters and to implement them in the model. Data from high spatial resolution optical sensors such as Sentinel-2 with a 10-m resolution are very promising and some products have recently came out such as Wang et al. (2020) on the lakes of the Tibetan plateau, or Zolfaghari et al. (2017) on the Erie Lake.

Lake area and depth are key parameters of the model. Thanks to satellite altimetry, water level changes as well as surface areas can be monitored. Future instruments as the next SWOT mission will soon provide these data at the global scale for lakes larger than 250 m x 250 m with a temporal resolution varying along the swath but better than 21 days (Biancamaria et al., 2016). Such data will be of tremendous interest to improve the calibration of both lake depth and area in FLake. Besides, the data will be also key to constrain water volumes, since resolving the lakes mass budget is the next development planned in our model.

Our preliminary results have shown that lakes are expected to impact the near-surface atmospheric fields. These impacts have already been shown by numerous studies such as the one of Le Moigne et al. (2016). The coupling of ORCHIDEE-FLake to the atmospheric model LMDZ is another mid-term perspective that will require developments in LMDZ to account for different atmospheric columns within an ORCHIDEE grid cell. But before, simulating the thermal processes in lakes provides the possibility to model other biogeochemical cycles, in particular GHG fluxes like carbon dioxide and methane. Then, the tools developed are opening the path to studies on the contribution of lakes to global warming and impacts of climate and environmental changes.

Code and data availability. The ORCHIDEE version developed for this project is available at https://doi.org/10.5281/zenodo.6383273. The model outputs which are archived on the LSCE servers are available from the corresponding authors upon request. The observed data and meteorological forcing used in this study are openly accessible online. The HydroLAKES database can be downloaded at https://hydrosheds.org/page/hydrolakes. The Global Observatory of Lake Responses to Environmental Change (GloboLakes) can be found at http://dx.doi.org/10.5285/76a29c5b55204b66a40308fc2ba9cdb3 and the Global Lake and River Ice Phenology Database at https://doi.org/10.7265/N5W66HP8.

Appendix A: Performance metrics

Bennett et al. (2013) describes several methods to characterize model performances. A simple way is to use an indice such as the Root Mean Square Error (RMSE) which is widely used to measure the performance of a model and is defined by Equation A1.

$$RMSE = \sqrt{\frac{1}{N}\Sigma(T_{sim} - T_{obs})^2} \tag{A1}$$

To better understand the errors, we decomposed the MSE (the square of the RMSE) into three components: the squared bias (SB), the lack of correlation weighted by the standard deviation (LCS) and the squared difference between standard deviations (SDSD). after normalization, the sum of these three components is equal to unity. The demonstration of this decomposition is done with the method described by Kobayashi and Salam (2000). LCS evaluates if the model correctly reproduces the temporal pattern and SDSD if the model gives the correct temporal amplitude. We also used the signed squared bias (SSB) to analyze the sign of the biases and identify systematic errors.

$$
\begin{aligned}
LCS &= 2SD_{sim}SD_{obs}(1-r) \\
SDSD &= (SD_{sim} - SD_{obs})^2 \\
SB &= (\overline{T_{sim}} - \overline{T_{obs}})^2 \\
SBB &= \text{sign}(\overline{T_{sim}} - \overline{T_{obs}})\,SB
\end{aligned}
\tag{A2}
$$

with

$$
\begin{aligned}
SD_{sim} &= \sqrt{\frac{1}{N}\sum(T_{sim} - \overline{T_{sim}})^2} \\
SD_{obs} &= \sqrt{\frac{1}{N}\sum(T_{obs} - \overline{T_{obs}})^2} \\
r &= \frac{\frac{1}{N}\sum(T_{sim} - \overline{T_{sim}})(T_{obs} - \overline{T_{obs}})}{\sqrt{\frac{1}{N}\sum(T_{sim} - \overline{T_{sim}})^2}\sqrt{\frac{1}{N}\sum(T_{obs} - \overline{T_{obs}})^2}}
\end{aligned}
\tag{A3}
$$

In some studies, the errors between the simulations and the observations are evaluated after removing the bias. This metric is the unbiased RMSE ($RMSE_{unbiased}$) defined with Equation A4. We also used this metric in our study.

$$RMSE_{unbiased} = \sqrt{\frac{1}{N}\Sigma(T_{sim} - T_{obs} - \text{sign}(\overline{T_{sim}} - \overline{T_{obs}})\sqrt{SB})^2} \tag{A4}$$

Author contributions. AB and CO designed the research, AB made the model developments, CO provides supervision and funding acquisition, AB and CO analyzed the results and wrote the manuscript.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgements. This work was supported by the French National Space Agency (Center National d'Etudes Spatiales) through the TOSCA-SPAWET program in preparation to the SWOT space mission. CEA (Commissariat à l'Energie Atomique et aux énergies alternatives) is also aknowledged for funding part of A. Bernus PhD grant. The authors would like to thank Karine Pétrus for her preliminary works with FLake model in our team and Fabienne Maignan for her precious advices on coding issues.

625

References

Ahmadzadeh Kokya, T., Pejman, A. H., Mahin Abdollahzadeh, E., Ahmadzadeh Kokya, B., and Nazariha, N.: Evaluation of salt effects on some thermodynamic properties of Urmia Lake water. I, 5(2), 343-348., International Journal of Environmental Research, 5, 343–348, 2011.

630 Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T., and Potes, M.: On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model, Tellus A: Dynamic Meteorology and Oceanography, 64, 15 829, https://doi.org/10.3402/tellusa.v64i0.15829, 2012.

Bastviken, D., Cole, J., Pace, M., and Tranvik, L.: Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate: Lake Methane Emissions, Global Biogeochemical Cycles, 18, 2004.

635 Beaulieu, J. J., DelSontro, T., and Downing, J. A.: Eutrophication will increase methane emissions from lakes and impoundments during the 21st century, Nature Communications, 10, 1375, https://doi.org/10.1038/s41467-019-09100-5, 2019.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrology and Earth System Sciences, 21, 589–615, https://doi.org/10.5194/hess-21-589-2017, 2017.

640 Bennett, N. D., Croke, B. F., Guariso, G., Guillaume, J. H., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D., and Andreassian, V.: Characterising performance of environmental models, Environmental Modelling & Software, 40, 1–20, https://doi.org/10.1016/j.envsoft.2012.09.011, 2013.

Benson, B.: Global Lake and River Ice Phenology Database, https://doi.org/10.7265/N5W66HP8, 2002.

Bernus, A., Ottlé, C., and Raoult, N.: Variance based sensitivity analysis of FLake lake model for global land surface modeling, Journal of
645 Geophysical Research: Atmospheres, https://doi.org/10.1029/2019JD031928, 2021.

Biancamaria, S., Lettenmaier, D. P., and Pavelsky, T. M.: The SWOT Mission and Its Capabilities for Land Hydrology, Surveys in Geophysics, 37, 307–337, https://doi.org/10.1007/s10712-015-9346-y, 2016.

Bonan, G. B.: Sensitivity of a GCM Simulation to Inclusion of Inland Water Surfaces, Journal of Climate, 8, 2691–2704, https://doi.org/10.1175/1520-0442(1995)008<2691:SOAGST>2.0.CO;2, 1995.

650 Bontemps, S., Boettcher, M., Brockmann, C., Kirches, G., Lamarche, C., Radoux, J., Santoro, M., Vanbogaert, E., Wegmüller, U., Herold, M., Achard, F., Ramoino, F., Arino, O., and Defourny, P.: Multi-year global land cover mapping at 300 m and characterization for climate modelling: achievements of the Land Cover component of the ESA Climate Change Initiative, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-7/W3, 323–328, https://doi.org/10.5194/isprsarchives-XL-7-W3-323-2015, 2015.

655 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T.,
660 Luyssaert, S., Madec, G., Madeleine, J., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore,

A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002010, 2020.

Bowling, L. C. and Lettenmaier, D. P.: Modeling the Effects of Lakes and Wetlands on the Water Balance of Arctic Environments, Journal of Hydrometeorology, 11, 276–295, https://doi.org/10.1175/2009JHM1084.1, 2010.

Carrea, L. and Merchant, C. J.: GloboLakes: Lake Surface Water Temperature (LSWT) v4.0 (1995-2016), https://doi.org/10.5285/76A29C5B55204B66A40308FC2BA9CDB3, type: dataset, 2019.

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, E., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J., Ghattas, J., Grandpeix, J., Idelkadi, A., Mellul, L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., Wang, F., and Zhao, Y.: Improved Near-Surface Continental Climate in IPSL-CM6A-LR by Combined Evolutions of Atmospheric and Land Surface Physics, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002005, 2020.

Choulga, M., Kourzeneva, E., Balsamo, G., Boussetta, S., and Wedi, N.: Upgraded global mapping information for earth system modelling: an application to surface water depth at the ECMWF, Hydrology and Earth System Sciences, 23, 4051–4076, https://doi.org/10.5194/hess-23-4051-2019, 2019.

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., Duarte, C. M., Kortelainen, P., Downing, J. A., Middelburg, J. J., and Melack, J.: Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget, Ecosystems, 10, 172–185, https://doi.org/10.1007/s10021-006-9013-8, 2007.

de Rosnay, P.: Integrated parameterization of irrigation in the land surface model ORCHIDEE. Validation over Indian Peninsula, Geophysical Research Letters, 30, 1986, https://doi.org/10.1029/2003GL018024, 2003.

Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., McDowell, W. H., Kortelainen, P., Caraco, N. F., Melack, J. M., and Middelburg, J. J.: The global abundance and size distribution of lakes, ponds, and impoundments, Limnology and Oceanography, 51, 2388–2397, https://doi.org/10.4319/lo.2006.51.5.2388, 2006.

Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P., Mironov, D., and Schar, C.: An offline study of the impact of lakes on the performance of the ECMWF surface scheme, Boreal Environment Research, 15, 100–112, 2010.

ESA: Land cover CCI product user guide version 2, Tech. rep., ESA, maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf, 2017.

Garnaud, C., MacKay, M., and Fortin, V.: A One-Dimensional Lake Model in ECCC's Land Surface Prediction System, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002861, 2022.

Golaz, J., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H., Lin, W., Lipscomb, W. H., Ma, P., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, Journal of Advances in Modeling Earth Systems, 11, 2089–2129, https://doi.org/10.1029/2018MS001603, 2019.

Guinaldo, T., Munier, S., Le Moigne, P., Boone, A., Decharme, B., Choulga, M., and Leroux, D. J.: Parametrization of a lake water dynamics model MLake in the ISBA-CTRIP land surface system (SURFEX v8.1), Geoscientific Model Development, 14, 1309–1344, https://doi.org/10.5194/gmd-14-1309-2021, 2021.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset: UPDATED HIGH-RESOLUTION GRIDS OF MONTHLY CLIMATIC OBSERVATIONS, International Journal of Climatology, 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Hartley, A., MacBean, N., Georgievski, G., and Bontemps, S.: Uncertainty in plant functional type distributions and its impact on land surface models, Remote Sensing of Environment, 203, 71–89, https://doi.org/10.1016/j.rse.2017.07.037, 2017.

Heiskanen, J. J., Mammarella, I., Ojala, A., Stepanenko, V., Erkkila, K.-M., Miettinen, H., Sandström, H., Eugster, W., Lepparanta, M., Jarvinen, H., Vesala, T., and Nordbo, A.: Effects of water clarity on lake stratification and lake-atmosphere heat exchange, Journal of Geophysical Research: Atmospheres, 120, 7412–7428, https://doi.org/10.1002/2014JD022938, 2015.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hostetler, S. W. and Bartlein, P. J.: Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur Lake, Oregon, Water Resources Research, 26, 2603–2612, https://doi.org/10.1029/WR026i010p02603, 1990.

Huziy, O. and Sushama, L.: Impact of lake–river connectivity and interflow on the Canadian RCM simulated regional climate and hydrology for Northeast Canada, Climate Dynamics, 48, 709–725, https://doi.org/10.1007/s00382-016-3104-9, 2017.

Jacob, D. and Podzun, R.: Sensitivity studies with the regional climate model REMO, Meteorology and Atmospheric Physics, 63, 119–129, https://doi.org/10.1007/BF01025368, 1997.

Kobayashi, K. and Salam, M. U.: Comparing Simulated and Measured Values Using Mean Squared Deviation and its Components, Agronomy Journal, 92, 345–352, https://doi.org/10.2134/agronj2000.922345x, 2000.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, Journal of the Meteorological Society of Japan. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Krinner, G.: Impact of lakes and wetlands on boreal climate, Journal of Geophysical Research, 108, https://doi.org/10.1029/2002JD002597, 2003.

Le Moigne, P., Colin, J., and Decharme, B.: Impact of lake surface temperatures simulated by the FLake scheme in the CNRM-CM5 climate model, Tellus A: Dynamic Meteorology and Oceanography, 68, 31 274, https://doi.org/10.3402/tellusa.v68.31274, 2016.

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Contoux, C., Cozic, A., Cugnet, D., Dufresne, J., Éthé, C., Foujols, M., Ghattas, J., Hauglustaine, D., Hu, R., Kageyama, M., Khodri, M., Lebas, N., Levavasseur, G., Marchand, M., Ottlé, C., Peylin, P., Sima, A., Szopa, S., Thiéblemont, R., Vuichard, N., and Boucher, O.: Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001940, 2020.

MacCallum, S. N. and Merchant, C. J.: Surface water temperature observations of large lakes by optimal estimation, Canadian Journal of Remote Sensing, 38, 25–45, https://doi.org/10.5589/m12-010, 2012.

MacKay, M. D.: A Process-Oriented Small Lake Scheme for Coupled Climate Modeling Applications, Journal of Hydrometeorology, 13, 1911–1924, https://doi.org/10.1175/JHM-D-11-0116.1, 2012.

740 Madec, G., Bourdallé-Badie, R., Pierre-Antoine Bouttier, Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J., Rousset, C., Storkey, D., Storto, A., and Vancoppenolle, M.: NEMO ocean engine, https://doi.org/10.5281/zenodo.1464816, 2008.

Malkki, P. and Tamsalu, R. E.: Physical Features of the Baltic Sea., 252, finnish marine research, edn., 1985.

745 Martynov, A., Sushama, L., Laprise, R., Winger, K., and Dugas, B.: Interactive lakes in the Canadian Regional Climate Model, version 5: the role of lakes in the regional climate of North America, Tellus A: Dynamic Meteorology and Oceanography, 64, 16 226, https://doi.org/10.3402/tellusa.v64i0.16226, 2012.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O.: Estimating the volume and age of water stored in global lakes using a geo-statistical approach, Nature Communications, 7, 13 603, https://doi.org/10.1038/ncomms13603, 2016.

750 Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Phillipps, P., Stouffer, R. J., and Swenson, S.: An Enhanced Model of Land Water and Energy for Global Hydrologic and Earth-System Studies, Journal of Hydrometeorology, 15, 1739–1761, https://doi.org/10.1175/JHM-D-13-0162.1, 2014.

Mironov, D.: Parameterization of lakes in numerical weather prediction. Description of a lake model., Tech. Rep. 11, Deutscher Wetterdienst, 2008.

755 Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., and Terzhevik, A.: Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO, Boreal Environment Research, 15, 218–230, 2010.

Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., Koven, C., Levis, S., Li, F., Riley, W., Subin, Z., Swenson, S., Thornton, P., Bozbiyik, A., Fisher, R., Heald, C., Kluzek, E., Lamarque, J.-F., Lawrence, P., Leung, L., Lipscomb, W., Muszala, S., Ricciuto, D., Sacks, W., Sun, Y., Tang, J., and Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM), Tech. rep., 760 UCAR/NCAR, https://doi.org/10.5065/D6RR1W7M, 2013.

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake (v1): Fenno-Scandinavian climate and lakes, Geoscientific Model Development, 11, 1321–1342, https://doi.org/10.5194/gmd-11-1321-2018, 2018.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hage-765 mann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative, Geoscientific Model Development, 8, 2315–2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Rooney, G. G. and Bornemann, F. J.: The performance of FLake in the Met Office Unified Model, Tellus A: Dynamic Meteorology and Oceanography, 65, 21 363, https://doi.org/10.3402/tellusa.v65i0.21363, 2013.

770 Salgado, R. and Le Moigne, P.: Coupling of the Flake model to the surfex externalized surface model, Boreal Environ Research Publishing Board, 15, 231–244, 2010.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol,

M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth System Science Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

Seland, \., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Semmler, T., Cheng, B., Yang, Y., and Rontu, L.: Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models, Tellus A: Dynamic Meteorology and Oceanography, 64, 17 339, https://doi.org/10.3402/tellusa.v64i0.17339, 2012.

Sima, S., Ahmadalipour, A., and Tajrishy, M.: Mapping surface temperature in a hyper-saline lake and investigating the effect of temperature distribution on the lake evaporation, Remote Sensing of Environment, 136, 374–385, https://doi.org/10.1016/j.rse.2013.05.014, 2013.

Stepanenko, V., Goyette, S., Martinov, A., Perroud, M., Fang, X., and Mironov, D.: First steps of a lake model intercomparison project: LakeMIP, Boreal Environment Research Publishing Board, 15, 191–202, 2010.

Stepanenko, V. M., Martynov, A., Jöhnk, K. D., Subin, Z. M., Perroud, M., Fang, X., Beyrich, F., Mironov, D., and Goyette, S.: A one-dimensional model intercomparison study of thermal regime of a shallow, turbid midlatitude lake, Geoscientific Model Development, 6, 1337–1352, https://doi.org/10.5194/gmd-6-1337-2013, 2013.

Subin, Z. M., Riley, W. J., and Mironov, D.: An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1, Journal of Advances in Modeling Earth Systems, 4, M02 001, https://doi.org/10.1029/2011MS000072, 2012.

Tebbs, E., Remedios, J., Avery, S., and Harper, D.: Remote sensing the hydrological variability of Tanzania's Lake Natron, a vital Lesser Flamingo breeding site under threat, Ecohydrology & Hydrobiology, 13, 148–158, https://doi.org/10.1016/j.ecohyd.2013.02.002, 2013.

Van de Walle, J., Thiery, W., Brousse, O., Souverijns, N., Demuzere, M., and van Lipzig, N. P. M.: A convection-permitting model for the Lake Victoria Basin: evaluation and insight into the mesoscale versus synoptic atmospheric dynamics, Climate Dynamics, 54, 1779–1799, https://doi.org/10.1007/s00382-019-05088-2, 2020.

Verpoorter, C., Kutser, T., Seekell, D. A., and Tranvik, L. J.: A global inventory of lakes based on high-resolution satellite imagery, Geophysical Research Letters, 41, 6396–6402, https://doi.org/10.1002/2014GL060641, 2014.

Viovy, N.: CRUNCEP Version 7 - Atmospheric Forcing Data for the Community Land Model, https://doi.org/10.5065/PZ8F-F017, type: dataset, 2018.

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., and Peylin, P.: Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): multi-scale evaluation of gross primary production, Geoscientific Model Development, 12, 4751–4779, https://doi.org/10.5194/gmd-12-4751-2019, 2019.

Wang, S., Li, J., Zhang, B., Lee, Z., Spyrakos, E., Feng, L., Liu, C., Zhao, H., Wu, Y., Zhu, L., Jia, L., Wan, W., Zhang, F., Shen, Q., Tyler, A. N., and Zhang, X.: Changes of water clarity in large lakes and reservoirs across China observed from long-term MODIS, Remote Sensing of Environment, 247, 111 949, https://doi.org/10.1016/j.rse.2020.111949, 2020.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model: ORCHIDEE SNOW MODEL EVALUATION, Journal of Geophysical Research: Atmospheres, 118, 6064–6079, https://doi.org/10.1002/jgrd.50395, 2013.

Wang, W., Rinke, A., Moore, J. C., Ji, D., Cui, X., Peng, S., Lawrence, D. M., McGuire, A. D., Burke, E. J., Chen, X., Decharme, B., Koven, C., MacDougall, A., Saito, K., Zhang, W., Alkama, R., Bohn, T. J., Ciais, P., Delire, C., Gouttevin, I., Hajima, T., Krinner, G., Lettenmaier, D. P., Miller, P. A., Smith, B., Sueyoshi, T., and Sherstiukov, A. B.: Evaluation of air–soil temperature relationships simulated by land surface models during winter across the permafrost region, The Cryosphere, 10, 1721–1737, https://doi.org/10.5194/tc-10-1721-2016, 2016.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, Journal of Hydrometeorology, 12, 823–848, https://doi.org/10.1175/2011JHM1369.1, 2011.

Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R., Schaefer, K., Jacobson, A. R., Lu, C., Tian, H., Ricciuto, D. M., Cook, R. B., Mao, J., and Shi, X.: The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data, Geoscientific Model Development, 7, 2875–2893, https://doi.org/10.5194/gmd-7-2875-2014, 2014.

West, W. E., Creamer, K. P., and Jones, S. E.: Productivity and depth regulate lake contributions to atmospheric methane: Lake productivity fuels methane emissions, Limnology and Oceanography, 61, S51–S61, https://doi.org/10.1002/lno.10247, 2016.

Zavialov, P. O., Izhitskiy, A. S., Kirillin, G. B., Khan, V. M., Konovalov, B. V., Makkaveev, P. N., Pelevin, V. V., Rimskiy-Korsakov, N. A., Alymkulov, S. A., and Zhumaliev, K. M.: New profiling and mooring records help to assess variability of Lake Issyk-Kul and reveal unknown features of its thermohaline structure, Hydrology and Earth System Sciences, 22, 6279–6295, https://doi.org/10.5194/hess-22-6279-2018, 2018.

Zolfaghari, K., Duguay, C. R., and Kheyrollah Pour, H.: Satellite-derived light extinction coefficient and its impact on thermal structure simulations in a 1-D lake model, Hydrology and Earth System Sciences, 21, 377–391, https://doi.org/10.5194/hess-21-377-2017, 2017.