



**HAL**  
open science

# Modélisation du bilan d'énergie des lacs dans le modèle de surface continentale ORCHIDEE

Anthony Bernus

► **To cite this version:**

Anthony Bernus. Modélisation du bilan d'énergie des lacs dans le modèle de surface continentale ORCHIDEE. Milieux et Changements globaux. Université Paris-Saclay, 2022. Français. NNT : 2022UPASJ002 . tel-03601672

**HAL Id: tel-03601672**

**<https://theses.hal.science/tel-03601672v1>**

Submitted on 8 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Modélisation du bilan d'énergie des lacs dans le modèle de surface continentale ORCHIDEE

*Modeling lake energy balance in the ORCHIDEE land  
surface scheme*

## Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 129, Sciences de l'Environnement d'Ile-de-France  
(SEIF)

Spécialité de doctorat : Météorologie, océanographie, physique de  
l'environnement

Graduate School : Géosciences, climat, environnement et planètes

Référent : Université de Versailles – Saint-Quentin-en-Yvelines

Thèse préparée dans l'unité de recherche **LSCE (Université Paris  
Saclay, CNRS, CEA, UVSQ)**, sous la direction de **Catherine Ottlé**,  
directrice de recherche, LSCE

Thèse préparée et soutenue à Gif sur Yvette, le 26 janvier 2022,  
par

**Anthony BERNUS**

### Composition du jury

<b>Jan POLCHER</b> Directeur de recherche, LMD	Président du jury
<b>Gianpaolo BALSAMO</b> Directeur de recherche, ECMWF	Rapporteur et Examineur
<b>Frédéric FRAPPART</b> Directeur de recherche, LEGOS/GET	Rapporteur et Examineur
<b>Marielle SAUNOIS</b> Enseignante-Chercheuse, LSCE/UVSQ	Examinatrice
<b>Wim THIERY</b> Professeur, VUB	Examineur
<b>Catherine OTTLE</b> Directrice de recherche, LSCE	Directrice

**Titre :** Représentation des lacs dans le modèle de surface continentale ORCHIDEE.

**Mots clés :** Modélisation du climat, Lac, Télédétection, Analyse de sensibilité, ORCHIDEE, Bilan d'énergie

**Résumé :** Au regard de l'importance des lacs sur le climat et de leur rôle sociétal, il est pertinent de les implémenter dans les modèles de climats pour comprendre leur influence sur les différents cycles comme ceux de l'eau, du carbone, du méthane ou encore leur interaction avec la couche limite de l'atmosphère. Cette thèse vise à implémenter les lacs dans le modèle de surface continentale ORCHIDEE. Nous avons choisi le modèle à une dimension Flake pour simuler le bilan d'énergie des lacs. Dans une première partie de cette thèse, nous avons réalisé une analyse de sensibilité des différents paramètres sur les températures et flux de surface simulés par le modèle Flake. Nous avons mis en avant que les températures de surface des lacs et les flux de chaleur de surface sont d'abord sensibles à la profondeur. Cependant, selon l'intervalle de profondeur choisi (peu ou très profond), on remarque que l'influence de la profondeur est plus importante pour les lacs peu profonds. Le coefficient d'extinction qui mesure la pénétration du rayonnement solaire dans l'eau est lui aussi influent pour les lacs de faible profondeur. Dans la suite du travail, nous avons couplé ORCHIDEE avec

FLake et testé plusieurs stratégies d'agrégation des profondeurs à plusieurs résolutions (0.25 et 0.5 degrés). Nous avons utilisé la base de données GloboLakes pour valider les températures de surface des lacs et la base de données « Global Lake and River Ice Phenology Database » pour valider les périodes de gel. Nous avons utilisé 5 forçages atmosphériques différents et nous avons remarqué la forte sensibilité du couplage aux forçages choisis. La stratégie d'agrégation des profondeurs des lacs à l'échelle de la maille de calcul, quant à elle, influence peu les résultats. Le modèle couplé a finalement montré de bons résultats. Pour chaque série, la médiane des erreurs obtenues sur les températures varie entre 2.7 K et 3.2 K. Les périodes de gel sont en général surestimées. La médiane des erreurs sur la période de gel varie entre 20 et 41 jours selon le forçage utilisé. Des pistes d'amélioration du couplage sont envisagées comme la paramétrisation spatiale du coefficient d'extinction et l'évolution temporelle des profondeurs contraintes par les mesures satellitaires du futur instrument SWOT. Enfin, dans une perspective à plus long terme, il est envisagé de coupler le modèle ORCHIDEE-FLake avec le modèle atmosphérique LMDZ.

**Title :** Representation of lakes in the ORCHIDEE land surface model.

**Keywords :** Climate modeling, Lake, Remote sensing, Sensitivity analysis, ORCHIDEE, energy balance

**Abstract :** Given the importance of lakes on the climate and their socio-economic role, it is important to implement them in climate models to understand their influence in the various cycles such as water, carbon, methane or energy cycle and their interaction with the atmospheric boundary layer. This thesis aims to implement lakes in the ORCHIDEE continental surface model. We have chosen the one-dimensional Flake model which simulates the energy balance of lakes.

In a first part of this thesis, we have performed a sensitivity analysis of the different parameters of the Flake model. We show that lake surface temperatures and surface heat fluxes are mainly sensitive to depth. However, depending on the chosen depth interval (shallow or very deep), we notice that the influence of the depth is more important for the shallow lakes. The extinction coefficient which measures the penetration of solar radiation in water is also influential for shallow lakes.

In the following work, we coupled ORCHIDEE with FLake and test several depth aggregation strategies at several resolu-

tions (0.25 and 0.5 degrees). We use the GloboLakes database to validate lake surface temperatures and the Global Lake and River Ice Phenology Database to validate freezing periods. We used five different atmospheric forcings and noticed the strong sensitivity of the coupling to the chosen forcings. The aggregation strategy of the lake depths has a small influence on the results.

Finally, the coupled model shows good results. For each series, the median of the errors obtained on the temperatures varies between 2.7 K and 3.2 K. The freezing periods are generally overestimated. The median of the errors on the freezing period varies between 20 and 41 days depending on the forcing used.

Future works are considered to improve the coupling, such as the spatial parameterization of the extinction coefficient and the temporal evolution of the depths constrained by the satellite measurements of the future SWOT instrument. Finally, in a longer term perspective, it is envisaged to couple the ORCHIDEE-FLake model with the atmospheric model LMDZ.

## Remerciement

Je voudrais d'abord remercier Catherine Ottlé pour avoir encadré cette thèse et pour sa disponibilité. Je souhaite remercier le CEA et le CNES pour avoir financé cette thèse.

Je tiens également à exprimer ma gratitude aux membres de mon jury qui ont accepté d'évaluer mon travail et avec qui j'ai apprécié échanger pendant la soutenance.

Je tiens à remercier Agnès Ducharne, Jan Polcher et Patrick Le Moigne pour votre participation durant mes comités de thèse. Je voudrais aussi remercier Patrick et son équipe (Aaron, Delphine, Thibault, Simon) pour avoir organisé le workshop de Toulouse en 2019 et de m'avoir accueilli plusieurs fois à Météo France à l'occasion de plusieurs réunions.

J'ai rencontré plusieurs personnes importantes pour le déroulement de cette thèse dans l'équipe MOSAIC. Je voudrais d'abord remercier Fabienne Maignan pour sa gentillesse et son aide dans le code ORCHIDEE, Pascal Maugis pour ses belles discussions que l'on a eu ensemble, Cédric Bacour pour ses petits conseils utiles et aussi Nicolas Voivy et Nicolas Vuichard pour leur aide sur certains points d'ORCHIDEE. Je voudrais remercier Sylvie Charbit pour avoir été ma marraine et de m'avoir soutenu surtout à la fin.

Je voudrais remercier le personnel du service administratif et informatique et en particulier Cécile Tessier pour son accueil chaque matin et Sandy Rigaux pour m'avoir accompagné dans les démarches administratives pour intégrer le LSCE.

Pendant cette thèse, je me suis investi aussi dans la vie du laboratoire et j'ai notamment rejoint le groupe de travail "Empreinte écologique" composé de volontaires pour réfléchir sur des pistes de réduction des émissions des GES liés aux activités du laboratoire. J'ai rencontré de très belles personnes au sein de ce groupe. Je tiens aussi à remercier les camarades syndicaux du CEA et aussi du LSCE avec qui j'ai partagé le sens de la solidarité et du collectif. Ces rencontres ont été très enrichissantes et m'a permis de rencontrer des personnes dans des domaines très différents au sein du CEA. Au LSCE, nous nous sommes aussi organisés autour des « pique-niques précaires » regroupant certains thésards, post-docs et CDD où nous avons pu débattre de notre avenir dans le monde de la recherche.

Au sein du LSCE, j'ai pu me lier d'amitié avec de nombreuses personnes. Les gens restent assez peu de temps et c'est avec tristesse que l'on voit partir les personnes proches au fur et à mesure. Je voudrais commencer à remercier Nina pour avoir partagé mon premier bureau, pour tes conseils sur le terminal linux, ton aide en anglais et de m'avoir accueilli durant les différents "social wenesday". J'ai rencontré de nombreuses personnes : Fanny P., Marine, Eric, Joël, Albert, Carley, Matthew, Johannes, Alex, Daniel, Spyros, Anne-Sofie, Fabio, Victoria, Elise, Audrey, Florentin et plus tard, j'ai pu rencontrer Yann, Maureen, Rodrigo, Jaime, Jonathan, Rose, Sophie, Ming, Elisa, Camille A., Martin, Omar. J'ai partagé de bons moments avec vous pendant les pauses-café et ces social wenesday.

J'ai notamment rencontré Henry au tout début de cette aventure au LSCE et qui partira assez tôt après la fin de son stage, mais avec qui je suis resté très proche durant la thèse.

Je remercie Zun avec qui j'ai pu avoir de belles discussions scientifiques et qui m'a permis de garder la motivation à des moments difficiles et aussi pour ses petites leçons sur l'écriture chinoise.

J'ai rencontré Alexandre avec qui j'ai partagé beaucoup de temps, nous avons échangé sur plein de sujets notamment sur le jardinage. Il m'a montré ses instruments de mesure, les lidars, et m'a intégré aux sessions de touch rugby.

J'ai rencontré un peu plus tard des personnes tout autant importantes pour moi par exemple autour de la nouvelle machine à café du premier étage (Marie, Julia, Gaëlle, Tiphaine, Stella, Clément D., Clément D., Mathis, Cyril, Fanny L., Robin, Sophie M., Davide, Camille C., Camille C., Germain, Alizée C., Alizée D.).

Je voulais particulièrement remercier Julia et Marie pour leur soutien dans la dernière ligne droite et d'avoir terminé ensemble cette épreuve et à Alizée pour cette soirée avant le jour de la soutenance.

Je voulais souhaiter bon courage à mon frère Maxime s'il s'aventure dans la voie du doctorat et de bien s'entourer pour pouvoir finir d'une belle manière et garder le moral jusqu'à la fin.

Je voulais aussi saluer mes ami·e·s de longue date que je n'ai pas pu voir aussi souvent que je l'aurais voulu pendant ces trois ans et demi.

Je voudrais enfin remercier Paula, la personne avec qui je partage ma vie. Cette thèse a été assez éprouvante psychologiquement où je n'avais plus de confiance en moi et où j'ai égaré mon optimisme. Cependant, les week-ends et les soirées où j'étais avec toi, je me sentais de nouveau moi-même et ta présence me ressourçait. C'est grâce à toi que j'ai pu finir cette épreuve.

# Table des matières

<b>Table des figures</b>	<b>7</b>
<b>Liste des tableaux</b>	<b>8</b>
<b>Abbréviations et Nomenclature</b>	<b>9</b>
<b>Introduction</b>	<b>15</b>
Les lacs, une composante du système Terre . . . . .	15
Les interactions avec le climat . . . . .	18
Importance des lacs dans la modélisation du climat . . . . .	21
Objectifs de cette thèse . . . . .	23
<b>1 Représentation du bilan d'énergie d'un lac</b>	<b>27</b>
1.1 Le lac, une composante du bilan d'énergie terrestre . . . . .	27
1.2 Bilan d'énergie d'un lac . . . . .	30
1.3 Structure thermique d'un lac . . . . .	31
1.4 Processus physiques à l'intérieur d'un lac . . . . .	34
1.5 Classification des lacs . . . . .	40
1.6 Modélisation du bilan d'énergie d'un lac . . . . .	42
<b>2 Modélisation des lacs dans ORCHIDEE</b>	<b>47</b>
2.1 Présentation d'ORCHIDEE . . . . .	47
2.2 Les différentes surfaces considérées par ORCHIDEE . . . . .	48
2.3 Bilan d'énergie ORCHIDEE/SECHIBA . . . . .	49
2.4 Présentation de FLake . . . . .	53
<b>3 Analyse de sensibilité du modèle FLake</b>	<b>63</b>
3.1 Introduction . . . . .	63
3.2 Méthode . . . . .	63
3.3 Synthèse des résultats . . . . .	64
3.4 Article 1 . . . . .	64
<b>4 Paramétrisation des surfaces et des profondeurs des lacs</b>	<b>85</b>
4.1 Répartition des lacs à l'échelle globale . . . . .	85
4.2 Spatialisation des profondeurs . . . . .	88
<b>5 Evaluation du couplage ORCHIDEE FLake</b>	<b>97</b>
5.1 Introduction . . . . .	97
5.2 Méthode . . . . .	97
5.3 Synthèse des résultats . . . . .	98

5.4 Article 2 . . . . .	98
<b>Conclusion</b>	<b>135</b>
Synthèse . . . . .	135
Perspectives . . . . .	138
<b>Bibliographie</b>	<b>161</b>

## Table des figures

1	Schéma simplifié du cycle de l'eau (source <a href="https://www.saur-saint-martin.fr/eau-et-environnement/respecter-lenvironnement/decouvrez-le-cycle-de-leau">https://www.saur-saint-martin.fr/eau-et-environnement/respecter-lenvironnement/decouvrez-le-cycle-de-leau</a> ) . . .	17
2	Carte des tendances de l'augmentation des températures estivales entre 1985 et 2009 en degrés celsius par décennie. (source <a href="#">O'Reilly et al. (2015)</a> ) . . . . .	20
1.1	Spectre du rayonnement solaire au sommet de l'atmosphère et au niveau de la mer (source <a href="https://en.wikipedia.org/wiki/Solar_irradiance##/media/File:Solar_spectrum_en.svg">https://en.wikipedia.org/wiki/Solar_irradiance##/media/File:Solar_spectrum_en.svg</a> ). . . . .	28
1.2	Bilan d'énergie moyen de la Terre (inspiré de <a href="#">Stephens et al. (2012)</a> , les rayonnements sont exprimés en $W/m^2$ ) . . . . .	30
1.3	À gauche, la masse volumique en fonction de la température de l'eau pour plusieurs valeurs de salinité. La courbe noire représente la courbe de solidification de l'eau et la courbe en pointillé indique des maxima de masse volumique. À droite, la température du maximum de densité de l'eau en fonction de la salinité et de la pression. . . . .	32
1.4	Schéma des différentes couches thermiques d'un lac . . . . .	34
1.5	Évolution du coefficient de traînée en fonction de la vitesse du vent à 10 m selon plusieurs études (adapté <a href="#">Wüest and Lorke (2003)</a> ) . . . . .	35
1.6	Etat de la turbulence et du profil de vent dans les couches limites atmosphérique et du lac (adapté de <a href="#">Wüest and Lorke (2003)</a> ) . . . . .	37
1.7	Modes d'oscillation des ondes internes d'un lac (adapté de <a href="#">Boegman (2009)</a> ) . . . . .	38
1.8	Les quatre régimes d'un lac selon la force du vent et la force de stratification (adapté <a href="#">Fischer et al. (1979)</a> ) . . . . .	39
2.1	Schéma de l'implémentation des lacs dans ORCHIDEE . . . . .	53
2.2	Profil de température simulé par FLake (a) en eau libre et (b) pour un lac gelé (adapté de <a href="#">Mironov (2008)</a> ) . . . . .	54
4.1	Carte d'occupation des sols de l'année 2015 du produit ESA-CCI à la résolution de 300 m. Les couleurs présentes sont associées à un type de surface et sont décrites par la table 4.1.1	86
4.2	Schéma représentant le satellite SWOT en orbite (source <a href="#">Fu et al. (2012)</a> ) . . . . .	92
4.3	Carte représentant le nombre de lacs traités par point de grille à une résolution de $0.5^\circ$ . . .	94
4.4	Représentation des fractions des lacs à une résolution de $0.5^\circ$ en haut avec un zoom représenté par un rectangle rouge sur la carte globale. Le tableau représente les distributions des profondeurs en bleu et du nombre de lacs en rouge pour chaque pixel de la région considérée. Chaque ligne représente une longitude et chaque colonne une latitude. . . . .	95



## Liste des tableaux

1	Classification de la formation des lacs par <a href="#">Hutchinson (1967)</a> . . . . .	16
2	Liste des modèles de climat qui intègrent la modélisation thermique des lacs . . . . .	25
1.1	Classification thermique des lacs de Lewis . . . . .	41
1.2	Liste des différents modèles thermiques de lacs . . . . .	45
4.1	Classification du produit ESA-CCI landcover . . . . .	87
4.2	Équations du modèle statistique utilisé dans la base de données HydroLAKES . . . . .	89
4.3	Discretisation de la carte des profondeurs . . . . .	93

## Abbreviations

**CLA** Couche Limite Atmosphérique

**CLL** Couche Limite du Lac

**LAI** Indice Foliaire (Leaf Area Index)

**LSM** Modèle de surfaces continentales (Land surface model)

**PFT** Type de végétation (Plant Functional Type)

**RMSE** Racine de l'erreur quadratique moyenne (Root mean square error)

**TKE** Energie cinétique turbulente (Turbulence kinetic energy)

**WASL** Zone d'Action des Vagues (Wave Affected Surface Layer)



## Nomenclature

$\Gamma_{\lambda_s}$  Constante liée à la paramétrisation de la conductivité de la neige

$\Gamma_{\rho_s}$  Constante liée à la paramétrisation de la densité de la neige

$\alpha$  Albédo

$\alpha_T$  Coefficient d'expansion thermique

$\alpha_w$  Albédo de l'eau

$\beta$  Résistance à l'évapotranspiration du sol

$\epsilon$  Emissivité

$\kappa$  Constante universelle de Von Karman

$\kappa_k$  Coefficient d'extinction dans la canopée du  $PFT_k$

$\kappa_w$  Coefficient d'extinction de l'eau

$\lambda$  Conductivité thermique du sol

$\lambda_I$  conductivité thermique de la glace

$\lambda_{seff}$  Conductivité thermique efficace de la neige

$\lambda_{sn}$  Conductivité thermique de la neige

$\left(\frac{dM_S}{dt}\right)_a$  Accumulation neigeuse liée à la précipitation

$\Phi_1$  Fonction de Businger du profil de vitesse

$\Phi_2$  Fonction de Businger du profil de température

$\Phi_\Theta$  Profil de température normalisée du lac

$\Phi_I$  Profil de température normalisée dans la couche de glace

$\Phi_Q$  Profil du flux de chaleur turbulent normalisée dans le lac

$\Phi_{Sn}$  Profil de température normalisée dans la couche de neige

$\rho_a$  Densité de l'air

$\rho_i$  Masse volumique de la glace

$\rho_s$  Masse volumique de la neige

$\rho_w$  Masse volumique de l'eau

$\sigma$  Constante de Stefan Boltzmann

$\tau_{CLA}$  Contrainte de cisaillement turbulent atmosphérique

$\tau_{CLL}$  Contrainte de cisaillement turbulent du lac

$\zeta$  Profondeur normalisée du lac

$A$  Surface du lac

$B$  Flux de flottabilité

$C_D$  Coefficient de traînée  
 $c_i$  Capacité calorifique massique de la glace  
 $C_Q$  Coefficient de forme du profil de flux turbulent de la thermocline  
 $c_s$  Capacité calorifique massique de la neige  
 $C_T$  Coefficient de forme du profil de température de la thermocline  
 $C_T^{min}$  Valeur minimale du coefficient de forme du profil de température de la thermocline  
 $c_w$  Capacité calorifique massique de l'eau  
 $C_{c1}$  Constante 1 de la paramétrisation de l'évolution de la hauteur de la couche de mélange  
 $C_{c2}$  Constante 2 de la paramétrisation de l'évolution de la hauteur de la couche de mélange  
 $c_{eff}$  Constante dans la paramétrisation de la conductivité efficace de la neige  
 $c_{l1}$  Constante 1 du modèle de Louis  
 $c_{l2}$  Constante 2 du modèle de Louis  
 $c_{l3}$  Constante 3 du modèle de Louis  
 $C_{TT}$  Coefficient de forme lié à la double intégration du profil de température de la thermocline  
 $d$  Distance de déplacement du profil de vent  
 $D$  Profondeur moyenne du lac  
 $F$  Fetch du lac  
 $f_k$  Fraction du  $PFT_k$   
 $f_i$  Fraction du PFT  $i$  dans une maille ORCHIDEE  
 $g$  Constante gravitationnelle de la Terre  
 $G$  Flux de conduction dans le sol  
 $g'$  Constante de gravité réduite  
 $H_L$  Flux de chaleur latent  
 $H_S$  Flux de chaleur sensible  
 $H_I$  Hauteur de la couche de glace du lac  
 $h_{ML_e}$  Hauteur de la couche de mélange du lac à l'équilibre  
 $h_{ML}$  Hauteur de la couche de mélange du lac  
 $H_{Sn}$  Hauteur de la couche de neige du lac  
 $I(z)$  Rayonnement solaire incident dans le lac à la profondeur  $z$   
 $K$  Constante de la formulation du coefficient de traînée de Charnock  
 $L$  Longueur de Monin-Obukhov  
 $L_f$  Energie d'évaporation de l'eau  
 $LW_{\downarrow}$  Rayonnement terrestre reçu  
 $LW_{\uparrow}$  Rayonnement terrestre émis

$N_{lac}$  Nombre de type de lac dans la maille  
 $p$  Pression  
 $p_{ref}$  Pression de référence (souvent égale à 1 bar)  
 $Q$  Flux de chaleur turbulent  
 $Q^*$  Flux turbulent caractéristique lié à la convection dans le lac  
 $q_a$  Humidité relative de l'air au niveau du sol  
 $q_a$  Humidité relative de l'air à une certaine hauteur  
 $Q_b$  Flux de chaleur à l'interface entre le lac et les sédiments  
 $Q_h$  Flux de chaleur à l'interface de la couche de mélange et de la thermocline  
 $Q_s$  Flux de chaleur à la surface  
 $Q_w$  Flux de chaleur à l'interface de la glace et de l'eau  
 $R_n$  Rayonnement net  
 $Ri$  Nombre de Richardson  
 $S_{100}$  Pente moyenne autour du contour du lac d'un rayon de 100 m  
 $SW_{\downarrow}$  Rayonnement solaire incident  
 $SW_{\uparrow}$  Rayonnement solaire réfléchi  
 $T$  Température  
 $T^*$  Température caractéristique lié à la turbulence  
 $T_a$  Température de l'air  
 $T_b$  Température à l'interface entre les sédiments et le lac  
 $T_f$  Température de gel de l'eau  
 $T_I$  Température à la surface de la couche de glace  
 $T_S$  Température de surface du lac  
 $T_{pot}$  Température potentielle  
 $t_{rc}$  Temps caractéristique de l'évolution du profil de température de la thermocline  
 $t_{rh}$  Temps caractéristique de l'évolution de la hauteur de la couche de mélange  
 $T_{Sn}$  Température à la surface de la couche de neige  
 $U(z)$  Vitesse de vent à la hauteur  $z$   
 $u^*$  Vitesse de friction de l'air  
 $U_{10}$  Vitesse de vent observée à 10 m  
 $W$  Nombre de Wedderburn  
 $w^*$  Vitesse de friction caractéristique de la turbulence  
 $z_0$  Rugosité du sol  
 $f_k^{max}$  Fraction du  $PFT_k$  sans la correction apportée par l'indice foliaire



# Introduction

## Les lacs, une composante du système Terre

Les lacs sont des étendues d'eau stagnante situées sur les surfaces continentales dont le renouvellement des eaux est très long. François Forel, un des pères fondateurs de la limnologie, décrivait les lacs ainsi dans son ouvrage *Der Seenkunde : allgemeine Limnologie* en 1901 : "On désigne par lac une masse d'eau stagnante sans communication directe avec la mer, située dans une dépression du sol fermée". Cette définition permet de séparer ces masses d'eau des rivières, fleuves et cours d'eau. L'origine du mot qui vient du latin "lacus" prenait déjà en compte le sens de dépression terrestre. Lacus voulait à la fois dire cuve, lac ou encore étendue d'eau stagnante. La création ou la disparition d'un lac se fait sur des temps géologiques assez longs. Cependant, selon les conditions et les causes de sa création ou de sa disparition, ces temps sont très variables. Par exemple, la formation du lac Baïkal, le plus vieux lac du monde, a commencé, il y a plus de 30 millions d'années après l'agrandissement d'une faille tectonique (Mats et al., 2011). Les grands lacs américains sont apparus quant à eux après la fonte de la glace de l'inlandsis laurentidien, il y a seulement 10 000 ans (Larson and Schaetzl, 2001). En règle générale, un lac peut émerger si une dépression terrestre existe avec un sol assez imperméable pour retenir l'eau, à condition que l'évaporation soit plus faible que l'apport en eau. L'origine de la dépression peut être diverse et variée. Hutchinson (1967) a créé une classification des lacs en fonction de leur origine. À cet effet, il proposa une classification avec 11 types de lacs (voir Table 1) divisée en 76 sous-types qui fait consensus dans la communauté limnologique.

L'origine des lacs peut expliquer certaines de leurs caractéristiques. Par exemple, les lacs sur une faille tectonique sont très profonds. Le lac Baïkal est un parfait exemple en la matière, avec une profondeur maximale atteignant 1 700 m. Le lac Tanganyika sur le grand rift africain est le second lac le plus profond (1 471 m). Ces lacs sont très allongés puisqu'ils suivent la forme d'une faille. Les lacs de type volcanique ou météorique ont l'aspect d'une cuvette bien caractéristique. Les lacs glaciaux sont quant à eux souvent très transparents. En effet, lorsque le glacier se retire et fond, les roches meubles sont érodées et les roches dures restent intactes ce qui induit des eaux très pures puisqu'elles contiennent peu de matières organiques et de sédiments. On peut aussi remarquer que ces lacs ont une bathymétrie particulière qui dépend de la manière dont fond le glacier. Pour plus de détails, Yao et al. (2018) a classifié et a décrit les lacs glaciaux, puis a expliqué la cause de leur forme en fonction de leur type.

Les lacs sont une composante à part entière du cycle de l'eau. La majeure partie de l'eau sur Terre est contenue dans les océans, à raison de 96 % du volume total, dont une fraction communique avec l'atmosphère et les surfaces continentales. Les



<b>Classes</b>	<b>Description</b>	<b>Exemples</b>
Tectonique	Bassin dans une faille	Lac Baikal (Russie) Lac Tanganyika (RDC, Tanzanie) Lac Vättern (Suède)
Volcanique	Cratère de volcan, barrage de coulée de lave	Lac Pavin (France) Lac Chūzenji (Japon)
Glissement de terrain	Retenue d'eau créée par un éboulis ou une coulée de boue	Lac de Usoi (Tadjikistan) Lac Waikaremoana (Nouvelle-Zélande)
Glaciaire	Lac créé par la fonte des glaces	Lac Erie (Etats-Unis) Lac Supérieur (Etats-Unis)
Karstique	Lac résultant de l'érosion hydrochimique de roches solubles	Lac Otjikoto (Namibie)
Fluviaux	Bras mort, lac de retenue, lac de delta, ...	Lac de Kanwar (Inde)
Éolien	Lac créé par déflation	Lacs dans le désert de Badain Jaran (Chine) ( <a href="#">Wang et al., 2016</a> )
Littoral	Flèche littorale, tombolo	Lac Pomorie (Bulgarie)
Organique	Lac créé par un barrage phytogène	Lac Ngaroto (Nouvelle-Zélande) Lac Rotokauri (Nouvelle-Zélande)
Anthropique	Retenue d'eau lié à un barrage anthropique	Lac Nasser (Egypte) Réservoir des Trois-Gorges (Chine)
Météorique	Lac à l'intérieur d'un cratère météorite	Lac Lonar (Inde) Lac Elgytgyn (Russie)

TABLE 1 – Classification de la formation des lacs par [Hutchinson \(1967\)](#)

lacs stockent seulement 0.8 % de l'eau terrestre sous forme liquide ([Messenger et al., 2016](#)). Ils recouvrent moins de 4 % des surfaces continentales ([Lehner and Döll,](#)



FIGURE 1 – Schéma simplifié du cycle de l'eau (source <https://www.sau-saint-martin.fr/eau-et-environnement/respecter-lenvironnement/decouvrez-le-cycle-de-leau>)

2004; Verpoorter et al., 2014) et sont répartis de façon inégale. La fonte des glaces de la précédente ère glaciaire dans l'hémisphère nord a laissé un grand nombre de lacs au Canada et en Scandinavie, ce qui explique une plus grande proportion de lacs dans ces endroits par rapport au reste du monde (Downing et al., 2006). Les lacs interagissent avec un grand nombre de composantes du cycle de l'eau. Froehlich et al. (2005) remarquent que ces interactions se font dans des proportions très diverses selon les lacs. Ils ont pu l'observer en mesurant des proportions d'isotopes comme ceux de l'oxygène. La figure 1 schématise les différentes interactions des lacs dans le cycle global de l'eau.

Grâce aux larges circulations d'air dans l'atmosphère, l'eau évaporée des océans peut ensuite être redistribuée sur les surfaces continentales via différents types de précipitations. L'eau a la possibilité d'être stockée en altitude sous forme de neige. Elle peut également s'infiltrer dans le sol ou s'écouler en surface. On définit alors le bassin versant, un domaine où l'eau converge à travers le ruissellement, les cours d'eau et les fleuves vers un même point d'arrivée, appelé exutoire. Celui-ci se situe pour les fleuves sur les côtes des mers ou des océans. Cependant, dans certains cas, l'eau n'a pas d'échappatoire et reste donc piégée dans les terres. Ainsi, un lac se forme, on le qualifie d'endoréique. La fonte des neiges alimente en partie ces cours d'eau. Les lacs, présents dans le bassin versant, interceptent une partie de cette eau et la stockent, ce qui module l'écoulement dans le bassin considéré (Spence, 2006). Par exemple, Bowling and Lettenmaier (2010) montre que 80 % de la neige fondue est interceptée par les lacs en Alaska. L'eau contenue dans les lacs peut se renouveler en moyenne au bout de quelques semaines jusqu'à

quelques années (Brooks et al., 2014). L'atmosphère a aussi des échanges directs avec les lacs. Les précipitations peuvent alimenter le lac et celui-ci peut évaporer en restituant de l'eau à l'atmosphère. Selon l'imperméabilité du sol, l'eau peut s'y infiltrer et être stockée dans des nappes phréatiques ou s'écouler à travers la roche jusqu'à l'océan. Des techniques comme la comparaison d'isotopes permettent de quantifier ce transfert d'eau entre les lacs et le sous-sol (Petermann et al., 2018). La proportion d'isotopes de l'oxygène, de l'hydrogène et du radon évolue différemment selon les milieux et les conditions météorologiques, ce qui permet d'estimer le temps de résidence de l'eau et donc les échanges avec les nappes souterraines ou le sous-sol.

## Les interactions avec le climat

Les échanges entre les lacs et l'atmosphère ne sont pas anodins. Pour illustrer cela, nous pouvons examiner l'exemple assez remarquable de la précipitation neigeuse au-dessus des grands lacs américains en hiver. L'air froid et sec qui se déplace au-dessus des grands lacs se charge en eau avec l'évaporation de ceux-ci relativement plus chauds. La colonne d'air est alors déstabilisée avec une température plus élevée à sa base et donne naissance à des mouvements convectifs. Des nuages se créent et se développent tout le long de leur trajet au-dessus des grands lacs. Quand ils quittent la zone d'influence du lac, la masse d'air se refroidit. Ainsi, l'humidité relative augmente jusqu'à atteindre le point de saturation. Cela se produit quand la température de l'air est en dessous du point de gel, des précipitations neigeuses apparaissent alors. Wright et al. (2013) ont étudié ce phénomène et montraient dans quelles conditions ces chutes neigeuses sont les plus importantes. Plus généralement, les lacs ont de nombreux effets sur les couches les plus basses de l'atmosphère. En effet, celles-ci répondent aux changements de propriétés de la surface sur des temps relativement courts, de l'ordre de l'heure. Au contraire, en excluant les cas de convections profondes, les couches de l'atmosphère plus hautes sont surtout influencées par des dynamiques de plus grande échelle comme celles engendrées par la force de Coriolis.

Comme on le constate avec l'exemple précédent, les lacs ont un réel impact sur le climat local. Le premier effet des lacs sur la Couche Limite Atmosphérique (CLA) est le lissage des températures de l'air. L'eau contenue dans les lacs a une capacité calorifique très élevée et le stockage de chaleur dans ceux-ci se fait sur une plus grande profondeur que pour d'autres types de surface. Ainsi, le lac absorbe le surplus d'énergie en diminuant les maxima de température et le restitue pendant les périodes plus froides (Bonan, 1995). Le lac étant une source importante d'humidité, il peut provoquer des événements de pluies plus fréquents et plus intenses. Par exemple, dans la région du lac Victoria en Afrique, les précipitations sont doublées la nuit par rapport aux régions environnantes (Van de Walle et al., 2020; Thiery et al., 2015). Les mouvements de circulation d'air sont grandement affectés par la

présence de lacs. [Chuang and Sousounis \(2003\)](#) ont pu montrer à l'aide de plusieurs simulations du modèle PSU-NCAR-MM5, comment les grands lacs américains influent sur la circulation de l'atmosphère.

L'état climatique de la Terre a de nombreux impacts sur les lacs qui sont sensibles à ces changements. [Adrian et al. \(2009\)](#) énumèrent quelles sont les variables clés des lacs qui sont affectées par le changement climatique et quantifient l'impact selon les différentes régions du monde. Ces changements imprègnent au cours du temps la couche de sédiments. Ainsi, les lacs peuvent archiver l'évolution du climat ([Leavitt et al., 2009](#)). Cette fonction est plus communément appelée intégrateur du système climatique. La cause principale du réchauffement climatique actuel est attribuée aux émissions anthropiques de gaz à effets de serre ([Friedlingstein et al., 2010](#); [Intergovernmental Panel on Climate Change, 2014](#)). Les lacs peuvent réguler le changement climatique en absorbant et stockant du  $CO_2$  dans les sédiments ([Cole et al., 2007](#); [Tranvik et al., 2009](#)). Dans d'autres cas, les lacs vont plutôt contribuer à émettre des gaz à effet de serre comme le méthane ([Bastviken et al., 2004](#)). L'émission de méthane non-anthropique provient entre 8 % et 16 % des lacs ([West et al., 2016](#)). Ces émissions peuvent évoluer avec le changement climatique comme celui-ci a un impact sur la structure thermique des lacs qui influence l'émission du méthane dans certains cas ([DelSontro et al., 2016](#); [Rasilo et al., 2015](#))).

Cette sensibilité climatique concède une certaine vulnérabilité des lacs au réchauffement global actuel, ce qui conduit par exemple à des problèmes d'approvisionnement en eau des êtres humains. Le changement d'état des lacs comme l'augmentation de la turbidité, de la salinité ou encore la fréquence des brassages qui perturbe la teneur en oxygène de l'eau peut affecter la faune et la flore. Dans certains cas, les lacs s'eutrophisent et permettent le développement de certaines algues ([Wagner and Adrian, 2009](#)). Ces conditions augmentent par exemple la mortalité piscicole. À court terme, la température de surface des lacs est la première variable affectée par le réchauffement climatique. Celle-ci a par exemple rapidement augmenté dans certains lacs alpins autrichiens où [Dokulil \(2014\)](#) mesurent une augmentation de  $0.6^{\circ}C$  par décennie, [Weckström et al. \(2016\)](#) mesurent un réchauffement de plus  $1^{\circ}C$  sur le mois de septembre pour certains lacs entre 1998 à 2009, [Woolway et al. \(2017\)](#) confirment ces tendances pour les lacs autrichiens.

Cette augmentation n'est pas uniforme sur Terre comme le montre la figure 2 provenant de [O'Reilly et al. \(2015\)](#). Sur un échantillon de 235 lacs, cette augmentation est en moyenne égale à  $0.34^{\circ}C$  par décennie entre 1985 et 2009 avec un intervalle entre  $-0.7^{\circ}C$  et  $1.3^{\circ}C$ .

L'augmentation globale des températures dépend non seulement du climat local, mais aussi de certaines caractéristiques du lac. Cette augmentation sur plusieurs décennies affecte ainsi les températures des eaux plus profondes ([Richardson et al., 2017](#)). Sur douze lacs observés, [Dokulil et al. \(2006\)](#) montrent que l'augmentation de température des eaux profondes de certains lacs européens est comprise entre

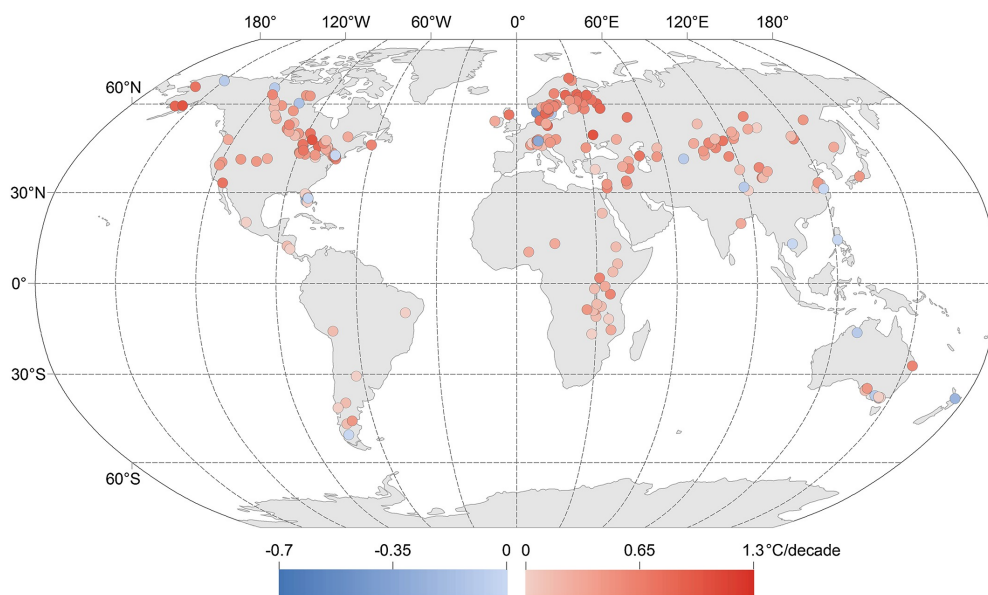


FIGURE 2 – Carte des tendances de l’augmentation des températures estivales entre 1985 et 2009 en degrés celsius par décennie. (source O’Reilly et al. (2015))

0.1°C et 0.2°C par décennie depuis le milieu du vingtième siècle. Le réchauffement climatique actuel a un réel impact sur les lacs sur le long terme comme en témoigne ce stockage de chaleur par les eaux profondes. Vanderkelen et al. (2020) ont estimé que les lacs et réservoirs ont stocké  $2.6 \pm .2 \cdot 10^{20} J$  d’énergie entre 1900 et 2020. Ce phénomène peut être amplifié, car les lacs gèlent moins longtemps avec le réchauffement climatique (Weckström et al., 2016; Thompson et al., 2009). En effet, pendant la période de gel, il y a plus de rayonnement réfléchi vers l’atmosphère que si l’eau du lac est libre. Il y a donc moins d’énergie stockée pendant les périodes de gel. Cela doit cependant être nuancé puisque la couche de glace sur le lac isole aussi de la perte de chaleur vers l’atmosphère. Ces changements de température en surface et en profondeur au cours des saisons affectent la stratification et les brassages internes des lacs (Piccolroaz et al., 2015; Zhong et al., 2016). Woolway and Merchant (2019) ont étudié la fréquence de ces brassages saisonniers sur 630 lacs répartis à travers le monde. Selon le scénario utilisé, à la fin du siècle, 60 à 100 lacs pourraient changer le nombre de brassage par an. Le changement de fréquence de ces brassages pourrait particulièrement affecter la distribution des nutriments et de gaz dissous comme l’oxygène dans la colonne d’eau (North et al., 2014; Yankova et al., 2017) et la couleur de l’eau à cause du brassage des sédiments (Andrew J. et al., 2008).

Enfin, le réchauffement climatique change même la distribution des lacs et leur forme. Zhang et al. (2019) ont recensé 300 lacs qui ont disparu dans la région du plateau mongolien en Chine et environ 100 lacs qui sont apparus dans l’Ouest

du pays entre 1960 et 2015. De même, l'Est des États-Unis voit son nombre de lacs augmenter contrairement à l'Ouest où celui-ci a tendance à diminuer pendant ces dernières décennies (Zou et al., 2018). Cette distribution change aussi dans les hautes altitudes. Salerno et al. (2014) ont montré la disparition de lacs dans les Alpes italiennes dans les basses altitudes. Cela est lié à une évaporation plus élevée avec l'augmentation des températures. Parallèlement, le nombre de lacs augmente dans les zones à plus haute altitude à cause de la fonte des glaciers. Cependant, ceux-ci sont voués à être éphémères avec l'amplification de l'augmentation des températures .

### Importance des lacs dans la modélisation du climat

Les lacs impactent le bilan d'énergie, le cycle de l'eau, les principaux cycles biochimiques, il semble donc pertinent de les prendre en compte dans les modèles. Un tel couplage permet aussi de comprendre l'impact du changement climatique sur les lacs. Dans les années 90, les modèles de climat avaient une résolution assez grossière d'environ 3° en latitude et en longitude pour la plupart. Ainsi, les lacs ne représentaient qu'une petite fraction quel que soit le point de grille considéré (inférieure à 30 %). Cependant, Pitman (1991) montra que le modèle de surface continentale BEST était sensible à l'ajout des lacs. Bonan (1995) a utilisé pour la première fois un modèle de climat (dans son étude, le modèle utilisé est celui du NCAR appelé CCM2) où il a implémenté le bilan d'énergie des lacs. Le modèle simule bien le lissage des températures de l'air et une évaporation plus forte dans les mailles où sont présents les lacs par rapport aux simulations sans lac.

Dans les années 2000, la communauté se focalise sur le couplage entre les modèles météorologiques et les modèles thermiques de lac pour comprendre les interactions entre l'atmosphère et ces derniers. Song et al. (2004) ont couplé le modèle régional NCSU-RegCM2 avec un modèle de lac à 3 dimensions paramétré pour le lac Victoria pour comprendre l'influence de celui-ci en Afrique sur le climat local. Long et al. (2007) ont couplé et confronté avec des observations le même modèle pour simuler le bilan d'énergie des lacs avec le modèle régional CRCM pour comprendre l'impact de deux grands lacs au Canada sur le climat local. Ils montrent aussi le lissage des températures et une plus grande évaporation des lacs en été. Dans le même temps avec le même modèle régional, Leon et al. (2007) ont aussi évalué l'impact du Grand lac des Esclaves (situé au Nord-Ouest du Canada) sur le climat avec le modèle CRCM, cependant celui-ci a été couplé avec un autre modèle de lac à 3 dimensions, ELCOM. En plus de montrer l'impact du lac sur l'évaporation, ils montrent le gradient de température de l'eau entre les bords et le centre du lac.

Progressivement, les lacs ont été implémentés dans des modèles de surfaces continentales (LSMs) pour prendre en compte l'impact qu'ils ont dans les simulations globales du climat. Dans certains cas, le lac est directement implémenté dans

la composante atmosphérique. En effet, son comportement physique diffère bien des surfaces végétales dans les LSMs et la représentation des surfaces se fait de manière composite dans les années 2000 (Molod, 2002), c'est-à-dire que les paramètres sont étalonnés en tenant compte de l'hétérogénéité des sols en ne résolvant qu'un seul bilan d'énergie. Ainsi, certains modèles ont implémenté un modèle de lac dans un module supplémentaire du modèle atmosphérique. Dans le cas du modèle de climat de l'IPSL, ce module a été introduit dans LMDZ (Krinner, 2003). L'approche "mosaïque" dans les LSMs (Molod, 2002) permet de représenter plus facilement d'autres types de surfaces comme les lacs ou encore les villes. Avec cette stratégie, plusieurs bilans d'énergie sont effectués par type de surface. Dans un premier temps, la représentation des lacs était très approximative. Par exemple dans le modèle de surface continentale SURFEX dans la version utilisée dans le modèle de climat CRNM-CM5.1, seuls les grands lacs étaient considérés et les plus petits (c'est-à-dire quand sa fraction est en dessous 50% de la maille) étaient remplacés par les autres surfaces de la même maille (Voldoire et al., 2013). De plus dans SURFEX, la température des grands lacs est supposée constante pendant la journée et elle est extrapolée par celle des points océans les plus proches chaque jour. Pour calculer les flux de chaleur au-dessus des grands lacs, les formules de Charnock ont été utilisées (Charnock, 1955). Aujourd'hui, les modèles de climat représentent les lacs avec des modèles plus représentatifs des processus présents dans les lacs. Par exemple, Dutra et al. (2010) a implémenté le modèle FLake qui simule le bilan d'énergie des lacs dans la composante continentale HTESSEL du modèle de climat de l'ECMWF. Le tableau 2 répertorie les différents modèles de climat qui ont implémenté un modèle de lac en leur sein. Ces derniers seront présentés dans le chapitre suivant.

Les modèles de lac utilisés dans les modèles de climat sont plutôt des modèles à une dimension ce qui permet d'optimiser le temps de calcul, mais ils sont aussi plus faciles à paramétrer. On peut remarquer que les deux modèles les plus utilisés sont le modèle FLake (Mironov, 2008) et le modèle LISSS (Subin et al., 2012) qui est une amélioration du modèle Hostetler avec une meilleure paramétrisation (Hostetler and Bartlein, 1990). Dans l'optique d'améliorer l'implémentation des lacs dans les modèles climatiques, certains modèles ont commencé à introduire le bilan de masse des lacs dans le cycle hydrologique de l'eau. Dans cette optique, Gronewold et al. (2020) ont développé le modèle L2SWBM pour fermer le bilan de masse des lacs dans l'optique de le coupler avec les systèmes de routage des LSM. Le modèle VIC est un modèle de surface continentale qui simule le bilan énergétique et le bilan hydrologique, et qui a été validé au niveau régional (MacCallum and Merchant, 2012) et global (Nijssen et al., 2001). Il modélise notamment le couplage des lacs avec le système de routage hydrologique qui a été implémenté par Cherkauer et al. (2003). Le modèle régional CRCM5 a implémenté la connexion entre les lacs et les rivières, ce qui a permis d'améliorer la simulation des débits dans les rivières (Huziy and Sushama, 2017). Certains modèles de surfaces continentales

ont introduit les lacs dans le cycle de l'eau pour des simulations climatiques globales. Dans un premier temps, cette implémentation était assez superficielle. Par exemple, le modèle CLM4 conserve la profondeur des lacs constante et la différence entre les précipitations et l'évaporation est directement ajoutée dans le terme de ruissellement qui va ensuite permettre d'alimenter les rivières du modèle (Oleson et al., 2013). Le modèle JSBACH adopte la même méthodologie. La première tentative d'introduire les lacs dans le routage des rivières dans un modèle de surface continentale en validant l'élévation de la surface d'eau de certains lacs fut réalisée par Milly et al. (2014). Récemment, un bilan de masse des lacs a été réalisé de manière complète dans le modèle de surface SURFEX et leur impact sur les débits de plusieurs bassins versants a été étudié (Guinaldo et al., 2021).

## Objectifs de cette thèse

Au regard de l'importance des lacs sur le climat et de leur rôle sociétal, il est pertinent de les implémenter dans les LSM pour comprendre leur comportement, mieux simuler les bilans hydrologiques et pouvoir prédire leur évolution future face aux changements globaux. Dans le cadre des développements du modèle de surfaces continentales ORCHIDEE, cette thèse vise à développer la représentation des lacs dans celui-ci. La première étape est l'implémentation du bilan d'énergie. Cette étape est la base pour de futurs développements comme l'intégration du bilan de masse des lacs dans le cycle de l'eau d'ORCHIDEE, ou encore l'intégration des lacs dans les cycles du carbone et du méthane. Les objectifs de cette thèse ont donc été les suivants :

- implémenter un modèle thermique de lac dans ORCHIDEE
- étalonner le modèle à l'aide de base de données globales
- évaluer les températures de surface et de période de gel simulées à l'aide d'observations

Pour répondre à ces objectifs, le manuscrit de thèse est décomposé en cinq chapitres. Le premier chapitre traite du bilan d'énergie des lacs. Nous y abordons le bilan d'énergie total de la Terre avant de nous concentrer sur le cas spécifique des lacs. Nous présentons aussi les processus physiques internes aux lacs pour appréhender leur bilan d'énergie. Dans le deuxième chapitre, nous présentons le modèle ORCHIDEE avec le module SECHIBA qui résout les bilans d'énergie et d'eau des surfaces continentales. Dans ce même chapitre, nous expliquons les développements réalisés pour introduire les lacs dans SECHIBA et plus particulièrement nous décrivons le modèle utilisé FLake. Dans le troisième chapitre, nous présentons une analyse de sensibilité de ce modèle qui nous a permis d'identifier quels sont les paramètres les plus influents du bilan d'énergie du lac. Le quatrième chapitre nous a permis d'explorer des pistes d'étalonnage de différents paramètres du modèle FLake en particulier de la profondeur dans la perspective du futur instrument SWOT (Surface Water Ocean Topography) qui permet de mesurer les



niveaux d'eau à chaque passage du satellite à l'aide d'un interféromètre. Un travail exploratoire d'assimilation des profondeurs dans le modèle FLake auquel nous avons contribué est présenté en Annexe 1. Grâce à l'analyse de sensibilité, nous avons constaté que la profondeur est le paramètre le plus important. Nous nous sommes donc focalisés sur celui-ci. Dans le cinquième chapitre, nous décrivons l'évaluation des températures simulées au niveau global. En conclusion, nous apportons des pistes d'amélioration du modèle et de son étalonnage, des pistes d'intégration des lacs dans les différents cycles pouvant être implémentés à la suite de cette thèse et des possibles prévisions climatiques réalisables avec ce nouveau couplage.

<b>Modèle climat</b>	<b>Composante continentale</b>	<b>Modèle thermique de lac</b>	<b>Référence implémentation</b>
BCC-CSM	BCC-AVIM (basé sur CLM)	Modèle LISSS (Subin et al., 2012)	Li et al. (2019)
CAMS-CSM	CLM	Modèle LISSS (Subin et al., 2012)	Rong et al. (2018)
CESM	CLM	Modèle LISSS (Subin et al., 2012)	Oleson et al. (2013)
CNRM	SURFEX	FLake (Mironov, 2008)	Salgado and Le Moigne (2010) Le Moigne et al. (2016)
CanESM5	CLASS	FLake (Mironov, 2008), Hostetler (Hostetler and Bartlein, 1990) ou le modèle CSLM (MacKay, 2012)	Martynov et al. (2012) MacKay et al. (2017)
E3SM	ELM (basé sur CLM)	Modèle LISSS (Subin et al., 2012)	Golaz et al. (2019)
EC-Earth3	HTESSEL	FLake (Mironov, 2008)	Dutra et al. (2010)
FGOALS	CLM4	Modèle LISSS (Subin et al., 2012)	He et al. (2019)
GFDL-CM4	LM4	Hostetler and Bartlein (1990)	Milly et al. (2014)
MPI-ESM	JSBASH	FLake (Mironov, 2008)	Reick et al. (2021)
NorESM2	CLM5	Modèle LISSS (Subin et al., 2012)	Seland et al. (2020)
SAM0-UNICON	CLM4	Modèle LISSS (Subin et al., 2012)	Park et al. (2019)
UKESM1-0-LL	JULES	FLake (Mironov, 2008)	Sellar et al. (2019) Rooney and Jones (2010)

TABLE 2 – Liste des modèles de climat qui intègrent la modélisation thermique des lacs



# 1 - Représentation du bilan d'énergie d'un lac

## Sommaire

---

<b>1.1</b>	<b>Le lac, une composante du bilan d'énergie terrestre . . . . .</b>	<b>27</b>
<b>1.2</b>	<b>Bilan d'énergie d'un lac . . . . .</b>	<b>30</b>
<b>1.3</b>	<b>Structure thermique d'un lac . . . . .</b>	<b>31</b>
1.3.1	Masse volumique de l'eau . . . . .	32
1.3.2	Nomenclature de la structure thermique . . .	33
<b>1.4</b>	<b>Processus physiques à l'intérieur d'un lac . .</b>	<b>34</b>
1.4.1	Mouvements liés au vent . . . . .	34
1.4.2	Stratification . . . . .	39
<b>1.5</b>	<b>Classification des lacs . . . . .</b>	<b>40</b>
<b>1.6</b>	<b>Modélisation du bilan d'énergie d'un lac . .</b>	<b>42</b>
1.6.1	Modèles 3D et 2D . . . . .	42
1.6.2	Modèle 1D . . . . .	42
1.6.3	Choix du modèle . . . . .	43

---

### 1.1 . Le lac, une composante du bilan d'énergie terrestre

L'objectif de cette thèse est l'intégration des lacs dans le modèle ORCHIDEE. Ainsi, avant de se focaliser sur les échanges d'énergie du lac avec son environnement proche, il est intéressant de définir le bilan d'énergie global de la Terre. Notre planète est le lieu de nombreux échanges d'énergie qui puisent leur source principalement du rayonnement solaire. La plus récente évaluation la fixe à  $1361.1 W \cdot m^{-2} \pm 0.5$  à l'entrée de l'atmosphère terrestre pour une surface exposée perpendiculairement aux rayons du Soleil par [Gueymard \(2018\)](#), cette valeur s'appelle la constante solaire. Celle-ci peut varier selon l'activité solaire (facule, tache solaire) ([Frohlich and Lean, 2004](#)). L'atmosphère n'est pas la seule à impacter le rayonnement reçu par la surface terrestre, l'inclinaison de celle-ci (dépendant de la latitude, du moment de la journée et de la saison) joue aussi un rôle. En moyennant cette énergie sur l'aire totale de la Terre, sa valeur est divisée par quatre, elle est environ égale à  $340.25 W \cdot m^{-2}$ . Cependant, le rayonnement solaire émis n'est pas constant et varie spectralement selon la longueur d'onde. La figure 1.1 représente l'éclairement solaire (aussi appelé irradiance) en fonction de la longueur d'onde. L'intégration spectrale de l'irradiance est égale à la constante solaire.

Un corps noir est par définition un corps qui absorbe la totalité du rayonnement reçu sans le réfléchir ou le transmettre. Il émet aussi un rayonnement avec un spectre continu qui suit la loi de Planck. Le rayonnement total émis est calculé à

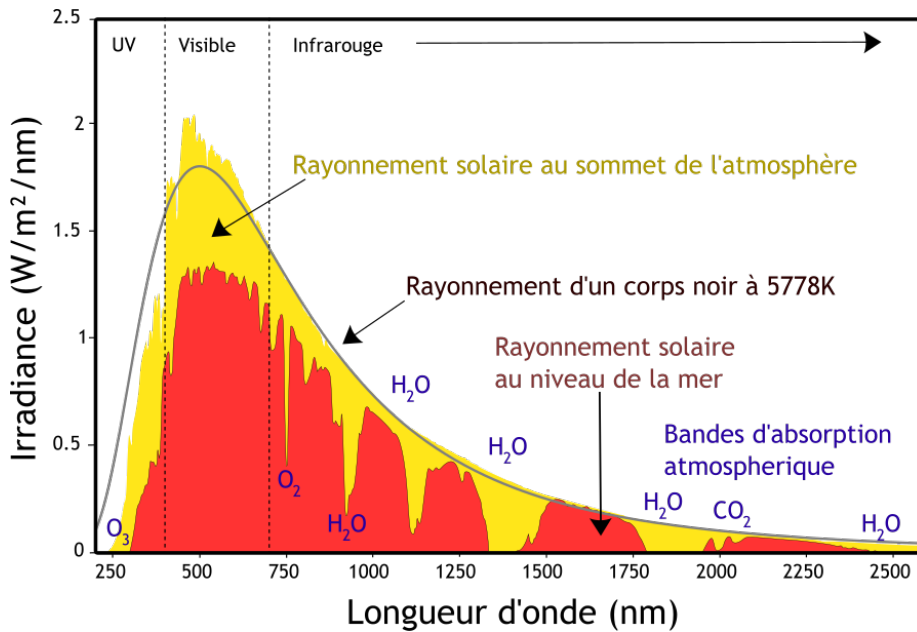


FIGURE 1.1 – Spectre du rayonnement solaire au sommet de l’atmosphère et au niveau de la mer (source [https://en.wikipedia.org/wiki/Solar\\_irradiance#/media/File:Solar\\_spectrum\\_en.svg](https://en.wikipedia.org/wiki/Solar_irradiance#/media/File:Solar_spectrum_en.svg)).

l’aide de la loi de Stefan-Boltzman et vaut  $\sigma T^4$  où  $\sigma$  est la constante de Stefan-Boltzmann et  $T$  est la température du corps noir.

La courbe en gris de la figure 1.1 représente l’éclairement d’un corps noir à la température de  $5778\text{ K}$  qui représente la température moyenne de la surface du Soleil. Cependant, l’éclairement solaire diffère à cause des raies d’absorption de l’atmosphère solaire. L’atmosphère terrestre, elle aussi, interagit avec le rayonnement solaire. En effet, des bandes caractéristiques sont observables. Sur la figure 1.1, on peut noter la présence de larges bandes d’absorption à  $900\text{ nm}$ , à  $1350\text{ nm}$  et à  $1850\text{ nm}$  qui sont caractéristiques des raies d’absorption de la vapeur d’eau dans l’atmosphère. Finalement, le rayonnement solaire moyen arrivant à la surface est d’environ  $165\text{ W/m}^2$  comme le montre la figure 1.2.

Une partie du rayonnement solaire est réfléchi par la surface terrestre. La grandeur qui mesure le rapport entre le rayonnement réfléchi et incident d’une surface s’appelle l’albédo. Celui-ci est une grandeur intégrale comprise entre 0 et 1. S’il est question de la fraction réfléchi pour une longueur d’onde donnée, on parle alors d’albédo spectral. Si on intègre l’albédo spectral sur le spectre solaire, sur l’ensemble de la surface terrestre sur une période assez grande, on peut estimer l’albédo moyen de la Terre à 0.303. Cependant, la surface terrestre est très hétérogène, l’océan a un albédo de 0.07 et la glace environ de 0.6. Ces dernières valeurs sont aussi obtenues à l’aide d’une intégration sur le spectre solaire. Avec ce petit exemple, on remarque que le réchauffement climatique fait fondre la glace qui est

remplacée par un sol avec un albédo plus faible. Le rayonnement absorbé est donc plus important ce qui contribue à renforcer le réchauffement climatique.

La Terre émet aussi de l'énergie sous forme de rayonnement appelé rayonnement terrestre. La Terre peut être assimilée à un corps gris qui est défini par une émissivité spectrale constante selon la longueur d'onde. L'émissivité, notée  $\epsilon$ , est la grandeur qui mesure le rapport entre le rayonnement émis et le rayonnement d'un corps noir porté à la même température pour une longueur d'onde donnée. Cette grandeur est inférieure à 1 pour les corps gris. Le rayonnement émis par la surface terrestre est estimé à  $390W/m^2$  comme le montre la figure 1.2. Avec la loi de Stefan, il est possible d'inverser une température terrestre à l'aide du rayonnement émis et de l'émissivité. Cette loi montre notamment que plus un corps noir est froid moins il émet de rayonnement, mais qu'il émet aussi plus de rayonnement électromagnétique dans les grandes longueurs d'onde. Ainsi, le spectre solaire est essentiellement contenu dans les ultraviolets et dans le spectre visible. À l'inverse, le spectre terrestre émet essentiellement dans la lumière infra rouge. Il y a très peu de recouvrement entre les deux spectres. L'atmosphère aussi émet son propre rayonnement avec une irradiance proche de celle de la surface terrestre. Celle-ci est émise vers l'espace, mais aussi vers la surface terrestre. La surface terrestre va absorber une partie de ce rayonnement. La grandeur qui mesure cela est l'absorptivité. La loi de Kirchoff postule qu'à l'équilibre thermique, l'absorptivité est égale à l'émissivité spectrale. Avec les différents concepts évoqués dans ce paragraphe, nous pouvons mettre en équation le bilan radiatif de la Terre avec l'équation 1.1 où  $R_n$  est le rayonnement net.

$$R_n = SW_{\downarrow} + SW_{\uparrow} + LW_{\downarrow} + LW_{\uparrow} \quad (1.1)$$

$SW_{\downarrow}$  est le rayonnement solaire qui atteint la surface terrestre.  $SW_{\uparrow}$  est le rayonnement solaire réfléchi. En définissant  $\alpha$  comme l'albédo de la surface, on obtient  $SW_{\uparrow} = \alpha SW_{\downarrow}$ .  $LW_{\downarrow}$  est le rayonnement terrestre réfléchi ou émis par l'atmosphère vers la surface.  $LW_{\uparrow}$  est la somme du rayonnement terrestre émis par la surface et du rayonnement atmosphérique réfléchi. Par convention, les flux sont négatifs quand ils sont dirigés vers l'atmosphère et positifs quand ils sont dirigés vers la surface terrestre. Finalement, on peut réécrire le bilan radiatif terrestre avec l'équation 1.2. Dans cette équation, toutes les variables sont moyennées sur la même surface.

$$R_n = (1 - \alpha) \cdot SW_{\downarrow} + LW_{\downarrow} - (\epsilon\sigma T_s^4 + (1 - \epsilon) \cdot LW_{\downarrow}) \quad (1.2)$$

Il faut compléter le bilan radiatif de la surface terrestre pour obtenir le bilan d'énergie total. Il faut ajouter le flux de conduction  $G$  qui permet de quantifier la chaleur absorbée par la surface de la Terre. En première approximation, la loi de Fourier permet de décrire ce flux en fonction de la profondeur. Ce flux est proportionnel au gradient de température, la constante multiplicative est la conductivité thermique exprimée en  $W/m/K$ . La surface terrestre peut aussi échanger de

l'énergie avec l'atmosphère sous forme de chaleur latente et/ou sensible, notée respectivement  $H_L$  et  $H_S$ . La chaleur latente est l'énergie transmise ou reçue après un changement d'état de l'eau. La chaleur sensible, quant à elle, est l'énergie échangée après un changement de température. Ces deux flux peuvent être conduits par des échanges turbulents, convectifs ou encore conductifs. En pratique, ces flux sont convectifs ou turbulents, la conduction thermique de l'air est trop faible et donc négligeable. Avec ces nouveaux termes, l'équation de base du bilan d'énergie à la surface de la Terre est la suivante :

$$(1 - \alpha) \cdot SW_{\downarrow} - (\epsilon\sigma T_s^4 - \epsilon \cdot LW_{\downarrow}) = G + H_S + H_L \quad (1.3)$$

Ces termes sont présentés sur la figure 1.2 avec une estimation de l'ordre de grandeur à l'échelle globale qui provient de [Stephens et al. \(2012\)](#).

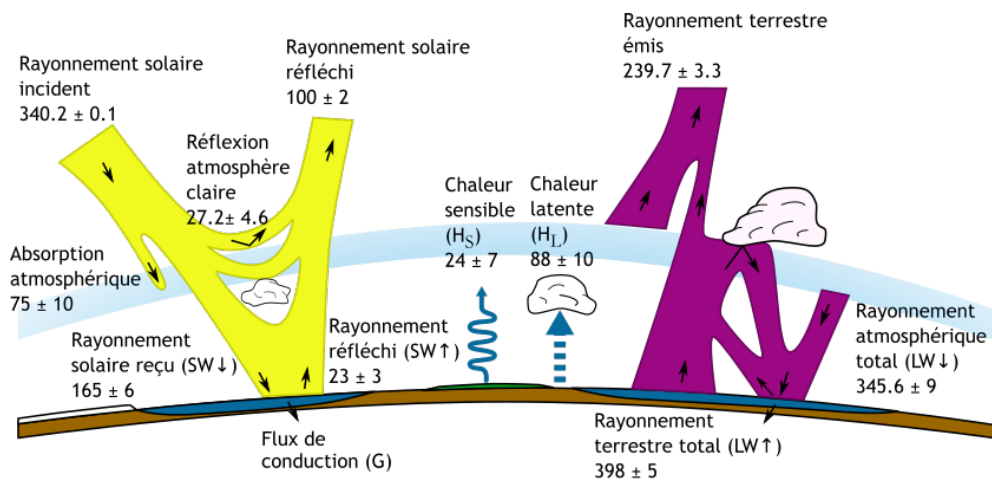


FIGURE 1.2 – Bilan d'énergie moyen de la Terre (inspiré de [Stephens et al. \(2012\)](#), les rayonnements sont exprimés en  $W/m^2$ )

## 1.2 . Bilan d'énergie d'un lac

Nous pouvons reprendre l'équation 1.3 dans le cas spécifique d'un lac. L'évaluation du bilan d'énergie à l'interface avec l'atmosphère dépend fortement de la surface considérée, car elle détermine l'albédo, l'émissivité ainsi que les différents flux de chaleur sensible et latente et le flux de chaleur pénétrant la surface du lac.

L'albédo de l'eau d'un lac est environ égal à 0.07 ([Lenters et al., 2005](#)) quand celui-ci n'est pas gelé. Cette valeur est largement en dessous de la moyenne terrestre de 0.3. Un lac absorbe donc une très grande partie du rayonnement solaire qu'il reçoit. Cependant, dès que les températures sont assez faibles notamment dans les hautes latitudes, les lacs peuvent geler. La glace se comporte totalement différemment et peut refléter jusqu'à 80 % du rayonnement solaire.

Le rayonnement émis dans les grandes longueurs d'onde par un lac suit aussi la loi de Stefan avec une émissivité très élevée et d'environ 0.99. À titre de comparaison, l'émissivité moyenne d'un sol nu est égale à 0.94 et est égale à 0.99 pour une forêt. Un lac émet donc plus de rayonnement infra-rouge que d'autres types de surface qui seraient à la même température.

L'eau est un milieu très transparent à la différence du sol où l'énergie absorbée est conduite à l'intérieur du milieu en suivant la loi de Fourier. Ainsi, le rayonnement solaire peut pénétrer dans le lac sur une grande profondeur. Ce transfert peut s'écrire selon la loi de Beer-Lambert (équation 1.4). Celle-ci permet de calculer le rayonnement transmis que l'on notera  $I(z)$  à une certaine profondeur  $z$  pour des milieux semi-transparents comme l'eau.

$$I(z) = (1 - \alpha_w)SW_{\downarrow}e^{-\kappa_w z} \quad (1.4)$$

Dans cette équation, on spécifie  $\alpha_w$ , l'albédo de l'eau. Le coefficient d'absorption, noté  $\kappa_w$ , est la grandeur physique qui mesure l'absorptivité du milieu. Plus ce coefficient est petit, plus le rayonnement pénètre dans le lac. Ce coefficient est spectral et varie selon la longueur d'onde du rayonnement incident. Cependant, dans une première approximation, nous pouvons construire un coefficient d'extinction moyenné sur le spectre solaire. Pour de l'eau pure, le coefficient d'extinction est égal environ à  $10^{-4} m^{-1}$  pour la lumière bleutée et à  $10^{-2} m^{-1}$  pour la lumière rouge. Ceci explique en partie pourquoi la mer paraît bleue sur les côtes, car l'eau est plus transparente pour cette couleur et donc le fond réfléchit seulement cette lumière. Cependant, l'eau dans les lacs n'est jamais pure, la présence de minéraux ou d'algues augmente le coefficient d'extinction. En général, quand il est moyenné spectralement, il vaut  $0.1 m^{-1}$  pour un lac clair et il peut monter jusqu'à  $3 m^{-1}$  pour des lacs les plus turbides.

L'équation 1.5 reprend le bilan d'énergie de l'équation 1.3 pour un lac. Le flux  $G$  est remplacé par l'intégrale sur la hauteur du lac de la loi de Beer-Lambert où  $D$  est la profondeur. Très souvent, le flux à l'interface du lac et de la couche de sédiments est très faible et il n'est donc pas pris en compte. Cependant, pour des lacs peu profonds, il faudrait prendre en compte les échanges thermiques avec les sédiments qui ne sont pas négligeables. Enfin, les flux de chaleur turbulents ne sont pas explicités ici et seront décrits par la suite.

$$(1 - \alpha_w) \cdot SW_{\downarrow} - (\epsilon_w \sigma T_s^4 - \epsilon_w \cdot LW_{\downarrow}) = \int_0^D (1 - \alpha_w) SW_{\downarrow} e^{-\kappa_w z} dz + H_s + H_L \quad (1.5)$$

### 1.3 . Structure thermique d'un lac

Pour estimer les flux de chaleur latente et sensible, il est important de connaître la température de surface du lac, le vent, la température et l'humidité relative de



l'air au-dessus du lac. La température de surface est conditionnée par les phénomènes de stratification thermique ou par des brassages qui sont expliqués par la force des vents ou par des inversions de densité.

### 1.3.1 . Masse volumique de l'eau

La masse volumique dépend à la fois de la température, de la pression et de la composition chimique comme la salinité. Pour un liquide, cette variation est très faible, mais suffisante pour créer certains mouvements convectifs qui vont modifier le profil de température. Le rapport de l'UNESCO (Millero et al., 1981) a fourni une équation précise pour calculer la masse volumique de l'eau sous une certaine température et une certaine salinité.

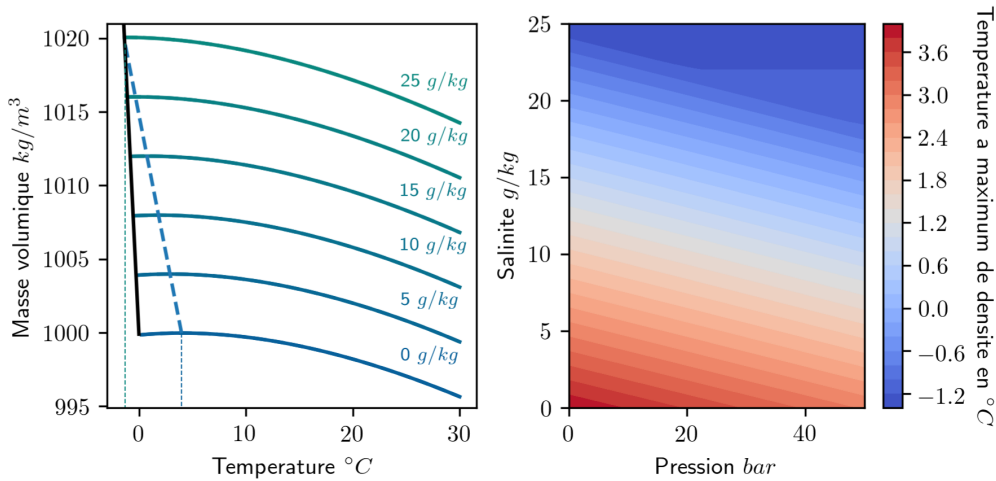


FIGURE 1.3 – À gauche, la masse volumique en fonction de la température de l'eau pour plusieurs valeurs de salinité. La courbe noire représente la courbe de solidification de l'eau et la courbe en pointillé indique des maxima de masse volumique. À droite, la température du maximum de densité de l'eau en fonction de la salinité et de la pression.

La figure 1.3 présente la masse volumique en fonction de la température, de la salinité et de la pression selon l'équation présente dans le rapport de l'UNESCO de 1981. La densité de l'eau diminue avec l'augmentation de la température. Cependant, une anomalie existe en dessous de  $4^{\circ}\text{C}$  pour l'eau pure (la courbe en bleu sur le graphe de gauche de la figure 1.3) et la densité augmente avec la température sous des conditions de pression normale ( $1\text{bar}$ ). Pour l'eau pure, la température du maximum de densité est environ égale à  $3.98^{\circ}\text{C}$ . À cette température, la densité vaut  $0.99997\text{ g/cm}^3$  sous une pression standard de  $1\text{bar}$ . Cette température de densité maximale diffère selon la salinité et la pression. Celle-ci se décale vers les températures plus faibles quand la salinité augmente jusqu'à croiser la température de gel pour une valeur de salinité de  $24.5\text{ g/kg}$ . L'effet de la pression agit de la

même manière sur la température de densité maximale comme on peut le voir sur le graphe de droite de la figure 1.3.

Dans les régions tempérées, cette anomalie de l'eau pure explique le brassage saisonnier après l'hiver. Ce brassage n'existe pas dans l'océan où la température maximale de densité est égale à la température de gel à cause de la forte salinité dans ce milieu-là environ égale à  $35\text{ g/kg}$ . D'autres phénomènes comme l'évaporation vont mélanger les eaux de surface dans l'océan. L'eau après évaporation voit sa salinité augmenter et donc aussi sa masse volumique, ce qui va créer un mouvement convectif. Dans les lacs profonds, la dépendance selon la pression doit aussi être prise en compte. En effet, le profil de température peut croiser la température de densité maximum puisque celle-ci diminue avec la pression. Quand cela se produit, on peut observer une instabilité appelée instabilité thermobarique. [Wüest et al. \(2005\)](#) décrivent ce phénomène et ses conséquences sur le lac Baïkal. Pour comprendre ce mécanisme, il faut utiliser la température potentielle et la densité potentielle. On définit la température potentielle par la température qu'aurait une particule fluide que l'on ramènerait à une pression standard  $p_{ref}$  de manière adiabatique (sans échange de chaleur avec l'extérieur). L'équation 1.6 permet de calculer cette température  $T_{pot}$  où sa dérivée par rapport à la pression  $p$  est souvent une expression compliquée.

$$T_{pot} = T + \int_p^{p_{ref}} \frac{\partial T}{\partial p} dp \quad (1.6)$$

La densité potentielle est la densité à la température potentielle. L'évolution de la densité potentielle permet de modéliser les instabilités thermobariques. En effet, si cette densité diminue avec la profondeur, la colonne d'eau est stable. Par contre, s'il y a une inversion de gradient de la densité potentielle, des mouvements convectifs apparaissent.

### 1.3.2 . Nomenclature de la structure thermique

Un lac est composé de plusieurs couches thermiques comme on peut le voir sur la figure 1.4. L'épilimnion, aussi appelé couche de mélange, est une zone où il y a une forte turbulence liée à la présence des vents en surface. La turbulence créée sous l'effet du vent homogénéise la température. En profondeur, l'eau est relativement calme et très peu d'énergie atteint les hautes profondeurs, ainsi la température reste constante et dans la plupart des cas, elle est égale à la température de densité maximale. Enfin une dernière couche, nommée métalimnion ou thermocline, permet de définir la zone de transition entre l'épilimnion et l'hypolimnion. Dans cette zone, la turbulence décroît fortement n'entraînant pas de mélange. Un fort gradient de température règne dans cette zone.

Selon les conditions météorologiques et la morphologie du lac, certaines couches peuvent disparaître. L'hypolimnion est absente des lacs peu profonds et la thermocline peut disparaître selon les saisons. Au contraire, dans certains cas de stratification, l'épilimnion peut disparaître. Dans d'autres cas plus complexes, plusieurs

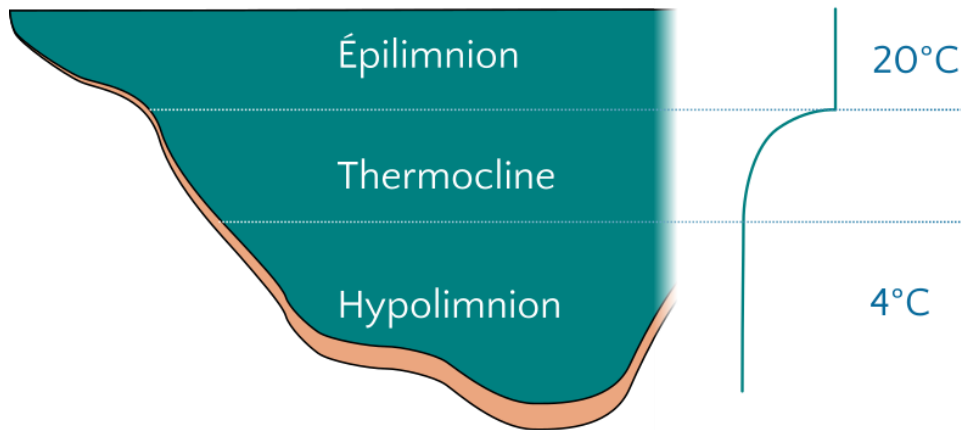


FIGURE 1.4 – Schéma des différentes couches thermiques d'un lac

thermoclines distinctes peuvent se succéder avec des couches homogènes sur la colonne d'eau.

## 1.4 . Processus physiques à l'intérieur d'un lac

### 1.4.1 . Mouvements liés au vent

Le vent est à l'origine de nombreux courants internes dans le lac, mais aussi de la turbulence qui règne dans la couche de mélange. Ainsi, connaître l'état de turbulence de l'atmosphère nous permet de remonter à l'information de la turbulence du lac. Par analogie avec CLA, la Couche Limite du Lac (CLL) est la zone directement influencée par le vent de surface.

### La contrainte turbulente dans l'atmosphère

La contrainte de cisaillement turbulent atmosphérique  $\tau_{CLA}$  va être transmise au lac, ce qui va permettre de générer la turbulence dans le lac. Elle est directement proportionnelle au carré de la vitesse du vent, mais elle dépend aussi de l'état de rugosité de la surface du lac. Cet état est caractérisé par le coefficient de traînée  $C_D$ . On exprime la contrainte de cisaillement  $\tau_{CLA}$  avec l'équation 1.7 où  $\rho_a$  est la densité de l'air et  $U(z)$  est la vitesse du vent à la hauteur  $z$ .

$$\tau_{CLA} = \rho_a C_D(z) U(z)^2 \quad (1.7)$$

$\tau_{CLA}$  dépend fortement de l'état de la surface et plus précisément des caractéristiques des vagues. Le vent influence directement  $\tau_{CLA}$ , mais il influence aussi

l'état de surface en créant les vagues. [Wüest and Lorke \(2003\)](#) propose d'étudier deux cas distincts selon la vitesse du vent. La figure 1.5 illustre ce propos avec des données issues de campagnes de mesures.

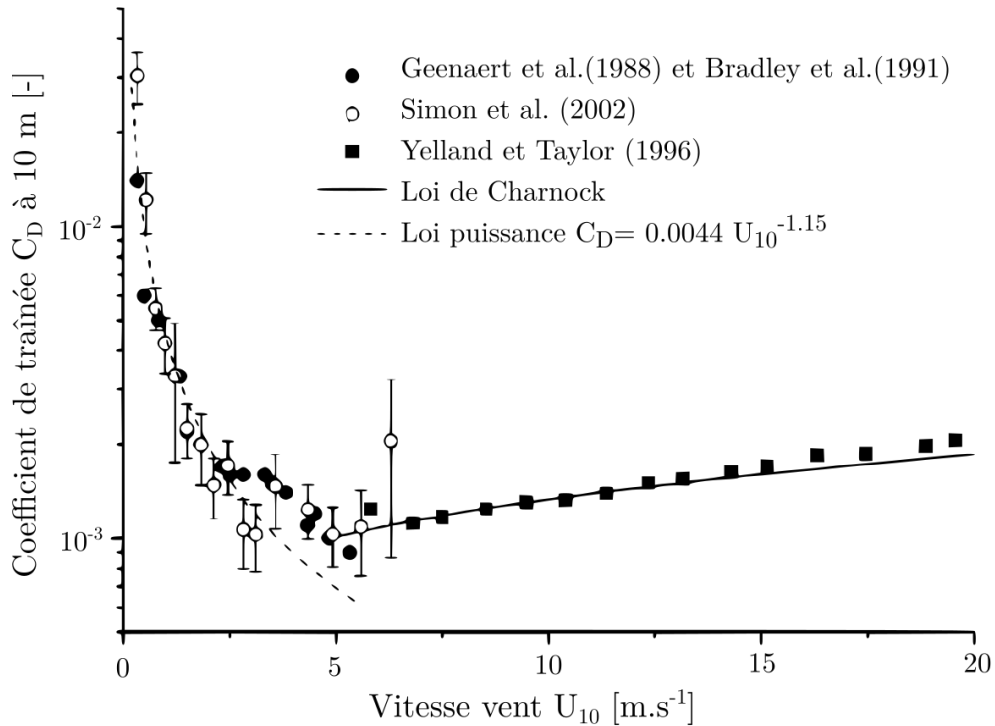


FIGURE 1.5 – Évolution du coefficient de traînée en fonction de la vitesse du vent à 10 m selon plusieurs études (adapté [Wüest and Lorke \(2003\)](#))

On observe un minimum du coefficient de traînée pour un vent de  $5 \text{ m s}^{-1}$ . En dessous de  $3 \text{ m s}^{-1}$ , ce coefficient est déterminé par la force de tension de surface ([Wu, 1994](#)). On peut ajuster une loi de puissance inverse entre  $C_D$  et  $U_{10}$  avec les observations des campagnes de [Bradley et al. \(1991\)](#) et de [Yelland and Taylor \(1996\)](#). L'équation de cette loi est donnée sur la figure 1.5.

Au-dessus de  $5 \text{ m s}^{-1}$ , la rugosité de la surface est conditionnée par la hauteur des vagues. [Charnock \(1955\)](#) formula une relation (voir équation 1.8) entre le profil de vitesse de vent et la vitesse de friction de l'air  $u^*$ .  $K$  est une constante établie expérimentalement à 12.5 et puis réévaluée à 11.3 par [Yelland and Taylor \(1996\)](#).  $\kappa$  est la constante universelle de Von Karman.

$$U(z) = u^* \left( \frac{1}{\kappa} \ln(gz/u^{*2}) + K \right) \quad (1.8)$$

$u^*/g$  est le paramètre qui est une longueur caractéristique de la hauteur des vagues avec  $g$  la constante gravitationnelle de la Terre, mais il est aussi propor-

tionnel à la rugosité de la surface qui permet de calculer  $C_D$ . Un dernier paramètre qui est l'âge des vagues joue un rôle sur le coefficient de traînée  $C_D$ . Terray et al. (1996) montre que les vagues plus jeunes font augmenter  $C_D$ . En effet, celles-ci n'ont pas encore atteint la vitesse maximale et sous l'effet de l'accélération elles déferlent et augmentent donc la rugosité.

On observe qu'au-dessus de  $5m/s$ ,  $C_D$  est conditionné par la hauteur des vagues (Stacey, 1999). La détermination de cette contrainte passe par la détermination du coefficient de traînée  $C_D$ .

La contrainte de cisaillement turbulente génère la turbulence dans la couche de mélange du lac. Sa transmission s'effectue de deux manières différentes : à travers les vagues (maintien du mouvement et du déferlement) et aussi par une contrainte de cisaillement turbulente  $\tau_{CLL}$  dans la CLL visible sur la figure 1.6. Ainsi, dans la Zone d'Action des Vagues (Wave Affected Surface Layer, WASL) juste en dessous de l'interface avec l'atmosphère, la contrainte turbulente est bien égale à  $\tau_{CLA}$  qui est décomposé par le terme lié au déferlement des vagues et à la contrainte turbulente dans le lac. Plus en profondeur la contrainte turbulente décroît, car le terme provenant des vagues disparaît pour atteindre la valeur de  $\tau_{CLL}$  comme le montre la figure 1.6. Cette diminution de la contrainte turbulente dans la CLL lié au déferlement des vagues est bien observée par Anis and Moum (1995), Buchard (2001) et Terray et al. (1996).

## La contrainte turbulente dans le lac

Nous avons décrit la transmission de la contrainte de la CLA à la CLL. Cependant, le vent transmet de l'énergie sous d'autres formes. Plus précisément, sous l'action du vent, la surface libre ou la thermocline s'incline emmagasinant de l'énergie potentielle comme sur la figure 1.7. Dès que le vent s'arrête, cette énergie se déplace sous forme d'onde, appelée seiche.

## Autres mouvements initiés par le vent

Les courants internes générés créent des cisaillements à l'intérieur du lac qui vont dissiper les seiches. Selon la force du vent, cette dissipation va prendre plusieurs formes. Spigel and Imberger (1980) établissent plusieurs régimes pour caractériser ces formes en utilisant le nombre de Wedderburn  $W$  (équation 1.9).

$$W = \frac{g' h_{ML}^2}{u^* F} \quad (1.9)$$

où  $g'$  est la constante de gravité réduite, c'est-à-dire la différence relative de la masse volumique entre l'épilimnion et la base de la thermocline multipliée par la constante gravitationnelle terrestre.  $h_{ML}$  est la hauteur de la couche de mélange.  $F$  représente le fetch qui est la longueur sur laquelle souffle le vent sans rencontrer

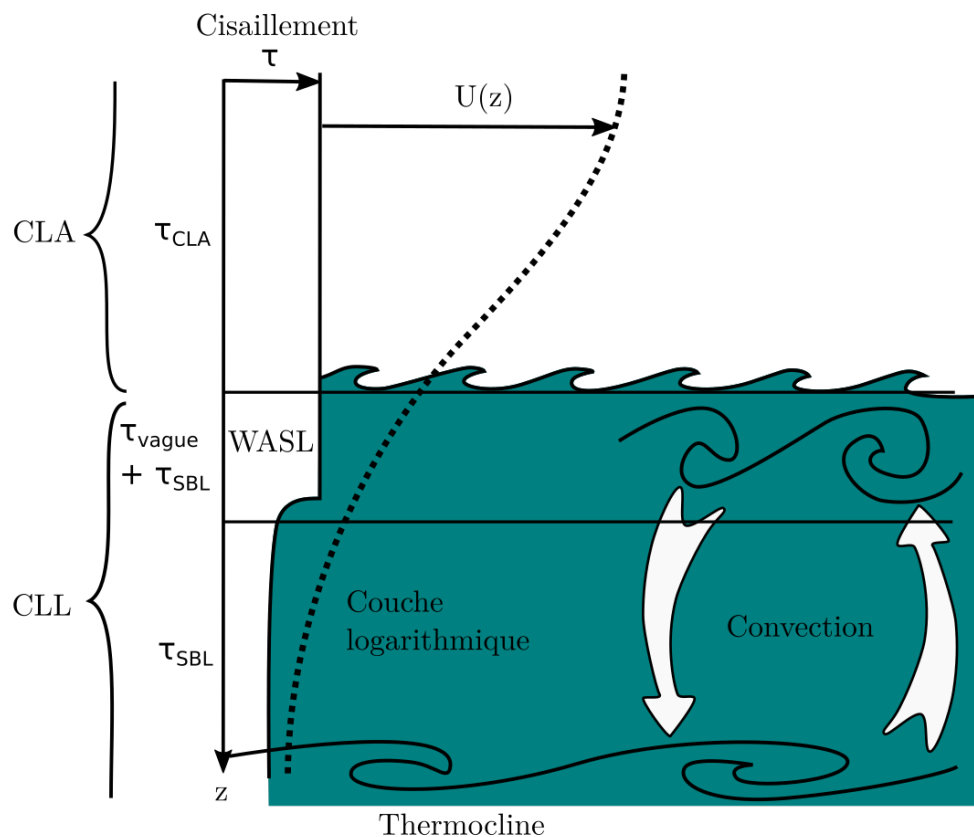


FIGURE 1.6 – Etat de la turbulence et du profil de vent dans les couches limites atmosphérique et du lac (adapté de [Wüest and Lorke \(2003\)](#))

d'obstacles, mesurée à partir de sa provenance (à partir de la berge dans notre cas).

Ce nombre représente le rapport entre la force liée à la turbulence du vent et la force stabilisatrice de flottabilité. Les régimes sont présentés par la figure 1.8. Les lignes correspondent à l'avancée du mécanisme au cours du temps et les colonnes au régime considéré. Si  $W$  est petit devant 1, alors la force du vent n'est pas suffisante pour créer des instabilités et il y aura seulement une inclinaison de la thermocline comme le régime A de la figure 1.8. Si  $W$  est grand devant 1, la thermocline va descendre au niveau de la berge créant de nombreuses instabilités qui vont mélanger une grande partie du lac, ce qui correspond au régime D. Enfin, quand  $W$  est environ égal à 1, la thermocline s'incline et des instabilités sont présentes créant de la turbulence sans toutefois mélanger complètement le lac. Ces instabilités peuvent être des paquets d'ondes internes dans le cas où  $W$  est légèrement inférieur à 1, ce qui correspond au régime B, ou des instabilités de type

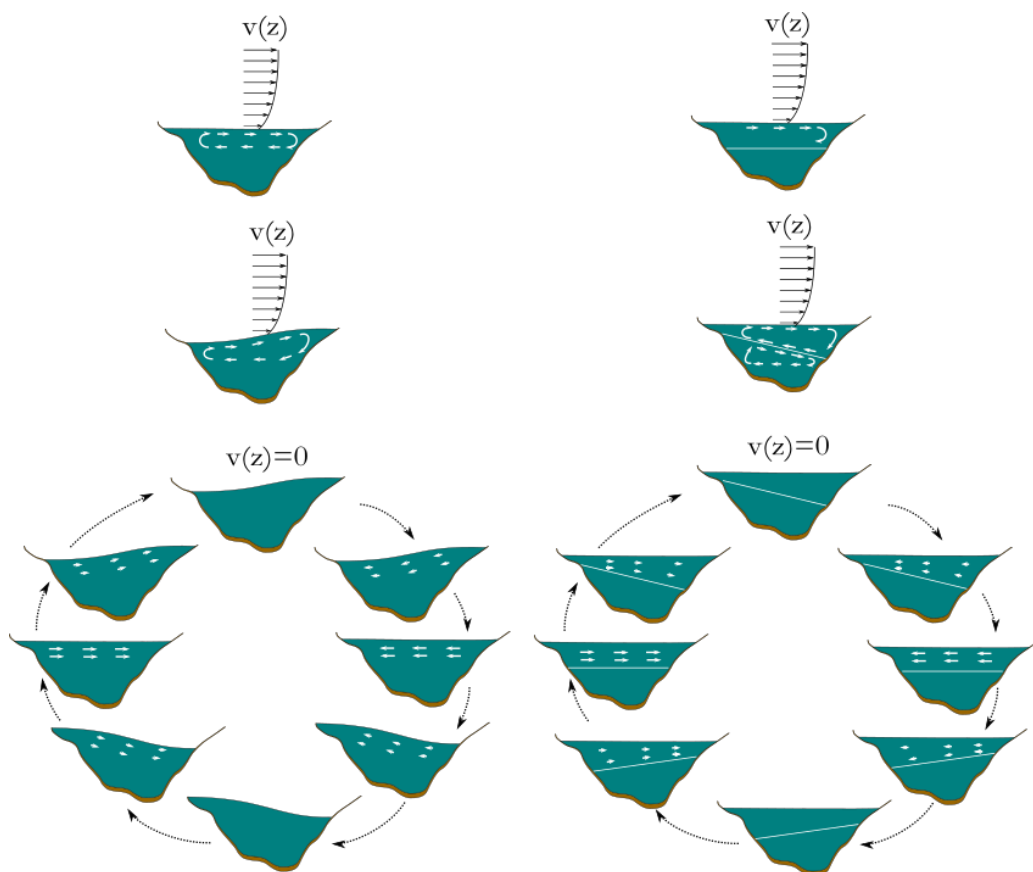


FIGURE 1.7 – Modes d'oscillation des ondes internes d'un lac (adapté de Boegman (2009))

Kelvin-Helmholtz quand  $W$  est légèrement supérieur à 1 pour le régime C.

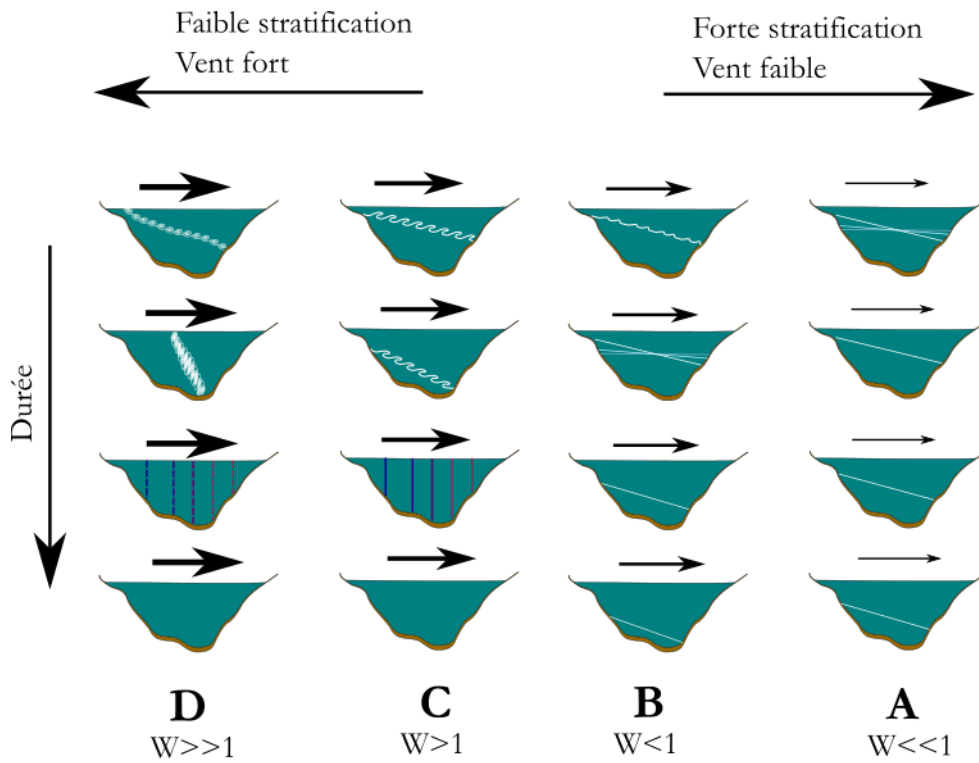


FIGURE 1.8 – Les quatre régimes d’un lac selon la force du vent et la force de stratification (adapté Fischer et al. (1979))

#### 1.4.2 . Stratification

Le processus en compétition avec la force du vent est la force de stabilisation liée au gradient de température. Pour quantifier ce phénomène, on définit le flux de flottabilité  $B$  (équation 1.10).

$$B = -\frac{g\alpha_T Q_s}{\rho c_p} \quad (1.10)$$

Ici, le produit  $g\alpha_T$  est le paramètre de flottabilité où  $g$  est la constante terrestre de gravité et  $\alpha_T$  est le coefficient d’expansion thermique. Pour l’air considéré comme un gaz parfait, le paramètre de flottabilité vaut  $g/T$ .  $Q_s$  est le flux d’énergie au niveau de la surface. Si le flux de flottabilité est positif, l’apport de chaleur tend à stratifier le lac, au contraire si le flux est négatif, alors des mouvements convectifs créeront de la turbulence qui sera caractérisée par une vitesse de friction  $w^*$  définie par l’équation 1.11

$$w^* = \sqrt[3]{h_{ML} \cdot B} \quad (1.11)$$



En été, dans les régions tempérées, on observe une forte stratification avec une couche de mélange très faible avec le fort rayonnement solaire. À la fin de l'été, on peut remarquer à certains endroits une alternance entre plusieurs couches homogènes et de thermoclines comme l'illustre l'étude de [Imberger \(1985\)](#). Le lac garde en mémoire la profonde stratification de l'été qui correspond à la thermocline la plus basse. Le vent vient mélanger la surface en créant une couche de mélange. Enfin une nouvelle stratification moins importante due à un rayonnement solaire plus faible de la fin de saison estivale vient restratifier le lac, mais dans une moindre mesure. On constate donc une alternance entre thermoclines et couches de mélange.

On retrouve cette stratification aussi au début de l'hiver juste avant que le lac gèle, l'énergie transmise à l'atmosphère permet de faire une stratification inverse ([Millero et al., 1981](#)). En effet, comme le montre la figure 1.3, la densité augmente avec la température en dessous de  $4^{\circ}\text{C}$ , ainsi dès que l'énergie est cédée à l'atmosphère, l'eau refroidit et se maintient en surface stabilisant ainsi la colonne d'eau. Si l'eau ne peut pas se refroidir davantage, elle se transforme en glace.

Dans les régions tempérées, on remarque souvent un brassage saisonnier au printemps et en automne. L'eau à la température de  $0^{\circ}\text{C}$  se réchauffe au printemps atteignant  $4^{\circ}\text{C}$ , celle-ci devient plus dense et plonge [Millero et al. \(1981\)](#). De même en automne le refroidissement rapide du lac au-dessus de  $4^{\circ}\text{C}$  tend à créer de grands courants convectifs. Ces phénomènes sont importants pour oxygéner le lac et maintenir une flore aquatique.

Selon les conditions limites liées à l'atmosphère et à la morphologie du lac, on a vu que l'on peut observer une forte stratification ou une grande couche de mélange. Cette compétition entre ces mécanismes permet donc le développement de la couche de mélange ou établit un gradient de température qui forme la thermocline.

## 1.5 . Classification des lacs

Les phénomènes présentés précédemment permettant de brasser ou de stratifier le lac dépendent du forçage. L'établissement d'une classification des lacs s'est construit pendant de très longues décennies. On peut remonter au premier essai de Geistbek en 1885 qui sépara les lacs chauds et froids en mesurant la température au fond du lac en été avec une limite de  $6.5^{\circ}\text{C}$ . La première classification mature fut celle de [Hutchinson and Loffler \(1956\)](#) en 1957. Elle sépare les plans d'eau en fonction de leur schéma de circulation. Cependant, un problème persistait avec la représentation des lacs peu profonds. [Lewis Jr. \(1983\)](#) proposa donc une subdivision du groupe des plans d'eau holomictiques pour compléter la classification de Hutchinson et Loffer. Un lac est holomictique s'il a la possibilité d'être brassé entièrement, au contraire les lacs méromictiques qui ne peuvent pas se mélanger sur la totalité de la colonne d'eau (lié par exemple à des barrières de densité ou

Type thermique	Sous-type	Propriétés Principales
Amictique		Totalement recouvert de neige toute l'année
Monomictique	Froid	Il y a un brassage complet une fois par an et seulement en été. Les températures restent proches de 4 °C
	Chaud	Il y a un brassage complet une fois par an et seulement en hiver
Dimictique		Il y a une stratification en hiver. Le lac est mélangé deux fois par an (au printemps et en hiver)
Polymictique	Froid discontinu	Le lac est gelé en hiver avec un brassage constant interrompu par quelques épisodes de stratification
	Froid continu	Le lac est gelé en hiver avec une stratification permanente interrompue par quelques épisodes de brassage
	Chaud discontinu	Le lac ne gèle pas et il est constamment mélangé avec quelques épisodes de stratification
	Chaud continu	Le lac ne gèle pas et il est stratifié mais il y a un brassage occasionnel de plus d'une fois par an
Oligomictique		Le brassage du lac ne se fait pas sur toute la colonne d'eau à cause d'une barrière chimique ou d'un hypolimnion très grand

TABLE 1.1 – Classification thermique des lacs de Lewis

chimique). Le caractère holomictique ou méromictique du lac dépend notamment de la profondeur du lac et ne détermine plus la catégorie du lac dans la nouvelle classification. La table 1.5 présente la classification de [Lewis Jr. \(1983\)](#).

## 1.6 . Modélisation du bilan d'énergie d'un lac

### 1.6.1 . Modèles 3D et 2D

Dans un premier temps, pour décrire le bilan d'énergie des lacs dans les modèles de climat (régionaux ou à l'échelle de la Terre), les modèles océaniques ont été utilisés et paramétrés dans le cas spécifique des grandes étendues d'eau continentale. En effet, les dynamiques entre les océans et les lacs sont assez similaires. On peut citer les modèles NEMO (Gurvan et al., 2017) et POM (Blumberg and Mellor, 1987) qui ont été utilisés par Song et al. (2004) et Dupont et al. (2012) pour simuler les lacs. Plus tard, des modèles à 3 dimensions ont été développés spécialement pour les lacs comme le modèle ELCOM (Hodges et al., 2000) ou encore Delft3D-Flow (Lesser et al., 2004) qui ont permis d'étudier l'interaction des lacs avec des climats locaux dans des modèles régionaux.

### 1.6.2 . Modèle 1D

Dans un souci de calcul numérique et de paramétrisation, des modèles à une dimension ont été développés pour être ensuite intégrés dans les modèles climatiques. Le premier concept utilisé a été assez primaire, il reposait sur une représentation du lac comme une seule couche de mélange. Par exemple les modèles de Ljungemyr et al. (1996) et Goyette et al. (2000) ont été développés avec ce concept.

Nous pouvons dénombrer deux familles de modèles. La première famille est fondée sur les méthodes de résolution des différences finies des équations des flux turbulents. La seconde est basée sur la résolution intégrale des équations sur la couche de mélange.

Les modèles de type intégral ont longtemps été délaissés quand il s'agissait de les inclure dans le modèle de climat. Néanmoins, MacKay (2012) a développé le modèle de type intégral CSLM. Certains auteurs antérieurs ont évoqué des soupçons sur leur paramétrage et leur stabilité quand ils sont couplés avec un modèle de climat. Un autre modèle de type intégral GLM (Hipsey et al., 2019) paramétrise l'énergie turbulente dans la couche de mélange lié à plusieurs processus (refroidissement, convection liée au gradient de densité, brassages liés au vent, brassages liés au cisaillement interne, turbulence de Kelvin-Helmoltz) et détermine après chaque processus le changement des températures dans la couche de mélange et sa hauteur. Dans une deuxième étape, le modèle calcule une diffusion turbulente dans l'hypolimnion avec une paramétrisation de l'énergie cinétique turbulente et de la dissipation turbulente en fonction de la fréquence de Brunt-Väisälä qui est la fréquence d'oscillation d'une particule fluide déplacée verticalement dans un environnement stable, ce qui permet d'actualiser la température dans la thermocline et l'hypolimnion. Des modèles plus simple comme Air2water Piccolroaz et al. (2013) repose sur la paramétrisation de la température et la hauteur de la couche de mélange avec la température de l'air.

Les modèles de type différentiel qui ont pour objectif d'être couplés avec un modèle de climat sont plus nombreux. Un premier groupe de modèles ferme l'équation de la chaleur en proposant une équation de transport pour l'énergie turbulente et de

dissipation turbulente. Ces modèles sont turbulents de type  $k - \epsilon$ . Les modèles de lacs qui reposent sur cette fermeture  $k - \epsilon$  sont les suivants : LAKEoneD (Joehnk and Umlauf, 2001), Simstrat (Goudsmit et al., 2002), LAKE (Stepanenko et al., 2016) et le modèle de Tsuang and Arpe (2001).

Au lieu de fermer le modèle avec deux équations, le coefficient de diffusion turbulente peut être paramétré plus simplement. Par exemple, Saloranta and Andersen (2007) développe un modèle dont le coefficient de diffusion turbulente est calculé avec le carré de la fréquence de Brunt-Väisälä. Cette idée pour fermer le terme de diffusion turbulente avec cette fréquence provient des travaux de Hondzo and Stefan (1993). Il est intéressant d'observer dans ce modèle que le rôle du vent est pris en compte dans une deuxième étape pour former la couche de mélange.

Le modèle de Hostetler and Bartlein (1990) est un modèle assez répandu dans les modèles de climat, il repose aussi sur une fermeture d'ordre 1 de l'équation de la diffusion turbulente. Celle-ci est réalisée à l'aide du nombre de Richardson (qui est paramétré par la fréquence de Brunt-Väisälä), du nombre de Prandtl et d'un paramètre qui représente la couche d'Ekman dans le lac. Cette paramétrisation se base sur les travaux de Henderson-Sellers (1985).

Le modèle LISSS est un modèle qui repose sur celui de Hostetler (Subin et al. (2012); Tan et al. (2015)) et qui améliore par exemple la paramétrisation de la turbulence. Ces modèles ont été couplés à des schémas de surface continentale, comme par exemple le modèle hydrologique VIC Liang et al. (1994) avec le modèle d'Hostetler.

Enfin, dans la troisième famille de modèle, nous retrouvons FLake développé par Mironov (2008). Le profil de température de la thermocline repose sur le concept de similarité du profil.

### 1.6.3 . Choix du modèle

La table 1.6.3 présente les différents modèles décrits ci-dessus. Dans notre étude, les modèles de lac en trois dimensions n'ont pas été retenus à cause de leur coût numérique élevé et du manque d'observations nécessaires pour l'étalonnage de leurs nombreux paramètres. Malgré tout, ces modèles ont été utilisés dans des simulations régionales sur des périodes assez courtes. Plusieurs inter-comparaisons de modèles à une dimension ont été réalisées pour guider le choix d'implémentation dans les modèles de climat. Ainsi, Martynov et al. (2010) ont comparé les deux principaux modèles thermiques de lacs utilisés dans les modèles de climat, c'est-à-dire FLake et le modèle Hostetler. Cette étude a conclu que FLake était meilleur pour simuler le bilan d'énergie sur les lacs américains même si les auteurs nuancent ces résultats, car des biais existent dans les forçages météorologiques. Stepanenko et al. (2013) ont fait une inter-comparaison avec 8 modèles de lacs à une dimension. Les modèles ont plutôt bien simulé les flux de chaleur et la température de surface. Malgré tout, FLake et le modèle de Hostetler ont plus difficilement simulé la structure thermique des lacs en été. Dans notre cas, nous avons fait le choix d'implémenter FLake dans ORCHIDEE pour son efficacité numérique et ses bonnes

performances. De plus, les paramètres internes au modèle ont été étalonnés pour tous les types de configuration, il est donc bien adapté pour être utilisé dans des simulations globales.

Modèle	Référence	Concept du modèle
NEMO	<a href="#">Dupont et al. (2012)</a>	3D Océan
POM	<a href="#">Song et al. (2004)</a>	3D Océan
GOTM	<a href="#">Sachse et al. (2014)</a>	1D Océan
LCOM	<a href="#">Hodges et al. (2000)</a>	3D
Delft3D-Flow	<a href="#">Lesser et al. (2004)</a>	3D
LAKEoneD	<a href="#">Joehnk and Umlauf (2001)</a>	1D Eq diff TKE fermeture d'ordre 2 $k - \epsilon$
Simstrat	<a href="#">Goudsmit et al. (2002)</a>	1D Eq diff TKE fermeture d'ordre 2 $k - \epsilon$
LAKE	<a href="#">Stepanenko et al. (2016)</a>	1D Eq diff TKE fermeture d'ordre 2 $k - \epsilon$
MyLake	<a href="#">Saloranta and Andersen (2007)</a>	1D Eq diff TKE fermeture d'ordre 1
Hostetler	<a href="#">Hostetler and Bartlein (1990)</a>	1D Eq diff TKE fermeture d'ordre 1
LISSS	<a href="#">Subin et al. (2012)</a>	1D Eq diff TKE fermeture d'ordre 1
VIC	<a href="#">Bowling and Lettenmaier (2010)</a>	1D Eq diff TKE fermeture d'ordre 1
ALBM	<a href="#">Tan et al. (2017)</a>	1D Eq diff TKE fermeture d'ordre 1
air2water	<a href="#">Piccolroaz et al. (2013)</a>	1D Forme intégrale
GLM	<a href="#">Hipsey et al. (2019)</a>	1D Forme intégrale TKE
DYRESM	<a href="#">Yeates and Imberger (2003)</a>	1D Forme intégrale TKE
CSLM	<a href="#">MacKay (2012)</a>	1D Forme intégrale TKE
FLake	<a href="#">Mironov (2008)</a>	1D Similarité du profil de température

TABLE 1.2 – Liste des différents modèles thermiques de lacs



## 2 - Modélisation des lacs dans ORCHIDEE

### Sommaire

---

<b>2.1</b>	<b>Présentation d'ORCHIDEE</b>	<b>47</b>
<b>2.2</b>	<b>Les différentes surfaces considérées par ORCHIDEE</b>	<b>48</b>
<b>2.3</b>	<b>Bilan d'énergie ORCHIDEE/SECHIBA</b>	<b>49</b>
2.3.1	Calcul du rayonnement net	49
2.3.2	Calcul des transferts de chaleur dans le sol	49
2.3.3	Calcul des flux de chaleur turbulents	49
2.3.4	Calcul des coefficients de traînée	50
2.3.5	Implémentation des lacs dans ORCHIDEE	52
<b>2.4</b>	<b>Présentation de FLake</b>	<b>53</b>
2.4.1	Concept de base	53
2.4.2	Le profil de température dans la thermocline	54
2.4.3	Les différents bilans d'énergie dans le lac	55
2.4.4	Calcul de la hauteur de la couche de mélange	57
2.4.5	Les couches de neige et de glace	58
2.4.6	Améliorations apportées par rapport à la version standard de FLake	60

---

### 2.1 . Présentation d'ORCHIDEE

ORCHIDEE est le modèle de surfaces continentales du modèle de climat de l'IPSL, il est donc couplé avec le modèle atmosphérique LMDZ mais aussi avec le modèle atmosphérique régional WRF dans la configuration REGIPSL. Certaines études s'intéressent à l'impact du climat sur les surfaces continentales, les conditions limites du modèle ORCHIDEE sont alors données par un forçage météorologique déjà prédéterminé et non par un modèle atmosphérique. Il simule plusieurs cycles comme celui de l'eau, du carbone et de l'énergie. Plus récemment le cycle de l'azote a été implémenté par [Vuichard et al. \(2019\)](#). Les différents cycles d'ORCHIDEE peuvent être simulés sur un lieu précis, sur une région donnée ou en global. Dans les deux derniers cas, la surface terrestre considérée doit être discrétisée sur une grille régulière où celle-ci est déterminée par la résolution du forçage atmosphérique utilisé ou par la résolution du modèle atmosphérique. Elle est typiquement de 2° pour le modèle de climat, mais peut descendre à quelques kilomètres pour des simulations régionales. [Laval et al. \(1981\)](#) développent dans une première version les équations du bilan d'énergie et de l'eau à la surface terrestre. La surface de la Terre va être différenciée en plusieurs types de surface végétalisée appelés Type de



végétation (Plant Functional Type) (PFT) (Ducoudré et al., 1993). Cette version s'appelle SECHIBA. Enfin, deux nouveaux modules ont fait leur apparition : STOMATE (Viovy, 1996) qui permet de décrire le cycle du carbone et LPJ (Sitch et al., 2003) qui décrit la dynamique de la végétation. Le modèle ORCHIDEE avec les trois composantes SECHIBA, STOMATE et LPJ est décrit pour la première fois par Krinner et al. (2005). Cependant, les 3 composantes ont des dynamiques sur différentes échelles de temps. Le cycle du carbone est actualisé tous les jours et la dynamique de la végétation, tous les ans alors que les cycles de l'eau et de l'énergie dans SECHIBA sont modélisés avec un pas de temps typique de 30 minutes.

## 2.2 . Les différentes surfaces considérées par ORCHIDEE

Dans SECHIBA, différents sols sont considérés pour représenter l'hétérogénéité de la surface. Jusqu'à aujourd'hui, seulement les sols végétalisés ou le sol nu sont pris en compte. Ainsi, dans les dernières versions d'ORCHIDEE, nous appelons les différents types de sol : PFT. La première version de SECHIBA (Ducoudré et al., 1993) comportait seulement 8 PFT. Puis, le nombre de PFT est monté à 13. Il existe dans cette version 8 types de PFT forêts qui sont caractérisées par le type de climat, de feuillage et sa persistance, 2 types de prairies et de cultures différenciées par leur mécanisme de photosynthèse C3 ou C4 et enfin le sol dénudé de végétation. Dans la dernière version d'ORCHIDEE (Lurton et al., 2020), il y a 15 PFT où un PFT prairie a été scindé en 3 pour représenter chaque type de climat (tropical, tempéré et boréal). La version utilisée dans ce travail est la version Tag 2.2 et elle comporte 15 PFT. Chaque PFT a ses propres caractéristiques biogéophysiques qui permettront de modéliser les différents cycles. Ainsi, une fraction de PFT est associée à chaque maille de la grille ORCHIDEE. Celle-ci représente la fraction de surface recouverte par le PFT sur la maille. Cependant, cette fraction est remplacée par une fraction effective calculée à partir de l'Indice Foliaire (Leaf Area Index) (LAI) pour pouvoir activer à chaque pas de temps la fraction effective de sol nu dans la maille. Cet indice représente le rapport entre la surface totale de feuilles projetées au sol et la surface considérée. Le LAI peut être simulé ou prescrit dans ORCHIDEE. Dans la version utilisée, la fraction effective est calculée à l'aide de l'équation 2.1 proposée par Monsi and Saeki (1953).

$$f_k = f_k^{max} (1 - e^{-lai\kappa_k})$$

$$f_1 = f_1^{max} + \sum_{k=2}^{Nb_{pft}} (f_k^{max} - f_k) \quad (2.1)$$

avec Coefficient d'extinction dans la canopée du  $PFT_k$  ( $\kappa_k$ ) le coefficient d'extinction dans la canopée du  $PFT_k$  occupant la fraction  $f_k$  de la maille.  $f_k^{max}$  est la fraction du  $PFT_k$  sans la correction apportée par le LAI.

## 2.3 . Bilan d'énergie ORCHIDEE/SECHIBA

Le module SECHIBA dans la version actuelle d'ORCHIDEE modélise un seul bilan d'énergie sur chaque maille. Un bilan hydrique est aussi modélisé où l'humidité du sol est calculée en fonction de la profondeur. Ce bilan se fait sur trois colonnes : une pour les forêts, une pour les prairies et les cultures et enfin une pour le sol nu. Le sol est discrétisé en 11 couches. Les phénomènes physiques suivants sont pris en compte pour simuler ce bilan : la précipitation, l'interception par la végétation, le ruissellement, l'extraction racinaire, le drainage, l'infiltration et enfin l'évaporation du sol nu et la transpiration de la végétation. Ce modèle est décrit par [de Rosnay \(2003\)](#).

### 2.3.1 . Calcul du rayonnement net

Le bilan d'énergie d'ORCHIDEE reprend l'équation 1.3. Les rayonnements solaire et terrestre provenant de l'atmosphère sont fournis par le forçage météorologique ou le modèle atmosphérique. L'albédo et l'émissivité de la surface sont les propriétés de la surface qui influencent le bilan radiatif. Dans ORCHIDEE, l'émissivité est approximée à 1 pour tous les types de surface et l'albédo est calculé en fonction de l'occupation des sols à l'aide de coefficients étalonnés à l'aide du produit albédo des satellites MODIS ([https://orchidas.lsce.ipsl.fr/dev/albedo/default\\_configuration.php](https://orchidas.lsce.ipsl.fr/dev/albedo/default_configuration.php)).

### 2.3.2 . Calcul des transferts de chaleur dans le sol

Le flux de chaleur dans le sol est modélisé à l'aide de la loi de Fourier. Le flux est directement proportionnel au gradient de température dont le coefficient de proportionnalité est la conductivité thermique du sol (équation 2.2) avec  $\lambda$  qui est la conductivité thermique du sol.

$$G = -\lambda \frac{\partial T}{\partial z} \quad (2.2)$$

Le sol est discrétisé par défaut en 18 couches selon une progression géométrique où la profondeur maximale est 90 m (sauf dans le cas d'un sol gelé). La condition limite de la dernière couche est un flux nul. En effet, le flux géothermique moyenné sur la Terre provenant de son centre est très faible et vaut 60 mW/m<sup>2</sup>.

### 2.3.3 . Calcul des flux de chaleur turbulents

Les flux de chaleur sensible  $H_S$  et latente  $H_L$  sont calculés avec l'équation 2.3. Ces flux sont modélisés par un processus de diffusion turbulente (très largement supérieure à la diffusion moléculaire).

$$\begin{aligned} H_S &= \rho_a |U(z)| C_D(z) C_p (T_s - T_a) \\ H_L &= L_f \beta \rho_a |U(z)| C_D(z) (q_s - q_{air}) \end{aligned} \quad (2.3)$$

On rappelle que  $T_S$  est la température de surface et  $C_D$ , le coefficient de traînée.  $\rho_a$  est la masse volumique de l'air.  $T_a$  est la température de l'air et  $q_a$  est l'humidité relative de l'air (en pratique mesurée à 2 m du sol).  $q_a$  est l'humidité relative de l'air au niveau du sol.  $L_f$  est l'énergie d'évaporation de l'eau.  $U(z)$  est la vitesse du vent (souvent mesurée ou fournie par la réanalyse à 10 m). La norme du vent est donnée par le forçage, mais la difficulté de ces équations est d'estimer le coefficient de traînée  $C_D$ . Les deux prochains paragraphes expliquent les théories utilisées pour le calculer. Dans le cas du flux de chaleur latente, il faut estimer le coefficient  $\beta$  qui exprime la résistance à l'évapotranspiration. Plusieurs coefficients  $\beta_k$  sont identifiés selon le processus physique considéré et le PFT. Les processus physiques pris en compte sont : la sublimation de la neige, l'interception et la transpiration de la canopée (différenciée selon le PFT), l'évaporation du sol nu, l'évaporation des plaines inondables. L'association de ces processus est faite de manière analogue à des circuits électriques de résistances montés en série et en parallèle.

### 2.3.4 . Calcul des coefficients de traînée

Dans notre cas, nous avons besoin de la description des profils de l'écoulement de l'air et de la température dans la CLA pour estimer correctement le coefficient de traînée. La théorie de Monin-Obukhov permet de relier ces profils avec les propriétés de la surface synthétisées à l'aide du coefficient de traînée  $C_D$ . Le coefficient de traînée dépend à la fois de la rugosité du sol nu  $z_0$ , de l'état de l'atmosphère et de la distance de déplacement  $d$  qui représente la hauteur où le profil de vent s'annule. L'état de l'atmosphère ici fait référence à sa stabilité. Celle-ci est directement liée au gradient de température potentielle. Quand il est positif, l'atmosphère est dite stable. En effet, une particule d'air qui serait déplacée dans un environnement plus chaud a tendance à revenir à son point de départ sous l'effet du gradient de température. À l'inverse, quand le gradient est négatif, le mouvement a tendance à être amplifié.

L'hypothèse principale de la théorie est de considérer les flux de contrainte de cisaillement turbulent, notés  $-\rho_a \overline{u'w'}$ , et le flux de chaleur turbulent  $Q$  constants dans la zone d'influence de la surface de la CLA. Ainsi, cette théorie reste valable dans les 50 premiers mètres, plus haut en altitude la force de Coriolis n'est plus négligeable.

Puisque la résolution directe des équations de Navier-Stokes n'est pas réalisable, cette théorie propose de généraliser le modèle de la longueur de mélange de Prandtl pour l'atmosphère stable et instable. Pour l'atmosphère neutre, sans gradient de température, le profil vertical du vent suit une loi logarithmique (voir équation 2.4) avec le modèle de mélange.

$$\frac{dU(z)}{dz} = \frac{u^*}{\kappa z} \quad (2.4)$$

La généralisation de ce profil utilise une analyse dimensionnelle basée sur le

théorème de Vaschy-Buckingham. Ce théorème énonce qu'à partir d'une équation à  $n$  variables avec  $k$  dimensions, il est possible d'obtenir une équation qui ne dépend plus que de  $n - k$  variables adimensionnelles. Pour les atmosphères non neutres, les variables d'intérêt sont  $u^*$ ,  $g/\overline{T_{pot}}$ ,  $\partial\overline{u}/\partial z$  (pour le profil de température  $\partial\overline{T}/\partial z$ ),  $z$ ,  $\overline{w'T'_{pot}}$  (le flux de chaleur sensible turbulent) qui sont exprimés avec trois dimensions, ainsi deux variables adimensionnelles  $\frac{\kappa z}{u^*} \frac{du(z)}{dz}$  (pour le profil de température  $\frac{\kappa z}{T^*} \frac{dT(z)}{dz}$ ) et  $z/L$  sont utilisées pour décrire le profil de vent (voir équation 2.5) où les fonctions  $\Phi_1$  et  $\Phi_2$  représentent les termes correctifs de l'équation 2.4. Pour le profil de température,  $T^*$  est une température caractéristique égale à  $-\overline{w'T'_{pot}}/u^*$ .  $L$  est une longueur caractéristique importante pour cette théorie qui permet de caractériser la stabilité de l'atmosphère qui se nomme longueur de Monin-Obukhov et est décrite par l'équation 2.6. Quand  $L$  est négatif, l'atmosphère est instable et inversement quand il est positif l'atmosphère est stable. L'atmosphère neutre est caractérisée par une valeur de  $L$  qui tend vers l'infini.

$$\begin{aligned} \frac{\kappa z}{u^*} \frac{dv(z)}{dz} &= \Phi_1(z/L) \\ \frac{\kappa z}{T^*} \frac{dT(z)}{dz} &= \Phi_2(z/L) \end{aligned} \quad (2.5)$$

$$L = \frac{u^{*3}}{\frac{g}{T} \overline{w'\theta'_v}} \quad (2.6)$$

$\Phi_1$  et  $\Phi_2$  sont fonction du paramètre  $z/L$  et ont été définies par [Businger et al. \(1971\)](#). Sans expliciter les détails techniques, les coefficients de traînée sont exprimés à l'aide de l'équation 2.7.

$$\begin{aligned} C_{D_{Temperature}} &= \left( \frac{\kappa}{\ln((z-d)/z_0) - \psi_1(z-d/L)} \right) \left( \frac{\kappa}{\ln((z-d)/z_{0t}) - \psi_2} \right) \\ C_{D_{vent}} &= \left( \frac{\kappa}{\ln((z-d)/z_0) - \psi_1(z/L)} \right) \end{aligned} \quad (2.7)$$

Une des difficultés est de calculer explicitement la longueur de Monin Obukhov  $L$  et la vitesse de friction  $u^*$ . L'idée générale est d'utiliser l'équation 2.5 du profil des vitesses avec une vitesse de vent observée à une certaine hauteur pour obtenir une valeur de  $u^*$  et  $T^*$  ( ou  $\overline{w'T'_{pot}}$ ) pour après obtenir  $L$ . Cependant, dans plusieurs cas, il n'est pas possible de le faire analytiquement. Des méthodes itératives créent par [Delsol et al. \(1971\)](#) ont été mises au point pour résoudre ce problème. ORCHIDEE utilise une autre méthode développée par [Louis \(1979\)](#). Elle propose une paramétrisation des fonctions de Businger  $\Phi_1$  et  $\Phi_2$  en fonction du nombre de Richardson  $Ri$  qui est décrit avec l'équation 2.8 au lieu de la longueur de Monin-Obukhov  $L$ .

$$Ri = \frac{g}{T_{pot}} \frac{\partial T_{pot}}{\partial z} \frac{1}{\left( \frac{\partial U}{\partial z} \right)^2} \quad (2.8)$$

L'idée du modèle repose sur un lien implicite entre  $Ri$  et  $L$ . Finalement, le coefficient de traînée est décrit par les équations 2.9 validées expérimentalement où  $c_{l1}$ ,  $c_{l2}$  et  $c_{l3}$  sont des constantes propres du modèle qui sont égales à 5. Pour rappel  $z_0$  est la rugosité du sol nu et  $d$  est la distance de déplacement.

$$\begin{aligned}
C_{d,neut}(z) &= \frac{\kappa^2}{\ln\left(\frac{z-d}{z_0}\right)\ln\left(\frac{z-d}{z_{0t}}\right)} \text{ pour } Ri = 0 \\
C_{d,stable}(z) &= \frac{C_{d,neut}(z)}{1 + 3c_{l1}Ri\sqrt{1 + c_{l3}Ri}} \text{ pour } Ri > 0 \\
C_{d,unstable}(z) &= C_{d,neut}(z) \frac{3c_{l1}Ri}{1 + 3c_{l1}c_{l2}C_{d,neut}(z)\sqrt{\frac{z-d}{z_0}Ri}} \text{ pour } Ri < 0
\end{aligned} \tag{2.9}$$

La dernière étape est d'estimer ce coefficient de traînée pour toute la maille en prenant en compte les différents PFT. Ceci est fait en moyennant les  $C_D$  calculés pour chaque PFT.

### 2.3.5 . Implémentation des lacs dans ORCHIDEE

Dans la version actuelle de SECHIBA, un seul bilan d'énergie est effectué, cependant les processus physiques des lacs sont très différents et nécessitent un bilan fait de manière indépendante. Ainsi, comme énoncé à la fin du chapitre précédent, le modèle FLake a été choisi pour faire ce bilan. Dans chaque maille, le rayonnement terrestre, les flux de chaleur sensible et latente de la surface lac et de la surface calculés par ORCHIDEE standard sont agrégés comme le montre la figure 2.1. De plus, selon la morphologie du lac et ses propriétés physiques, le lac réagit de manière différente. Nous avons donc donné la possibilité de construire plusieurs types de classe de lac au sein d'une même maille qui pourront différer par leurs paramètres. Le calcul des flux agrégés et de la température de surface de la maille est donné par l'équation 2.10.

$$\begin{aligned}
H_{S_{agg}} &= \left( \sum_{PFT_i} f_i \right) * H_{s_{ORCH}} + \sum_i^{N_{lac}} f_i H_{si} \\
H_{L_{agg}} &= \left( \sum_{PFT_i} f_i \right) * H_{L_{ORCH}} + \sum_i^{N_{lac}} f_i H_{Li} \\
T_{s_{agg}} &= \sqrt[4]{ \left( \sum_{PFT_i} f_i \right) * T_{s_{ORCH}}^4 + \sum_i^{N_{lac}} f_i T_{si}^4 }
\end{aligned} \tag{2.10}$$

L'indice *agg* représente les variables agrégées au niveau de la maille.  $N_{lac}$  est le nombre de type de lac dans la maille et  $f_i$  est la fraction du PFT  $i$  dans la maille. Pour simuler les différents bilans d'énergie, les différents paramètres de FLake peuvent être spatialisés. La version d'ORCHIDEE développée dans cette thèse ne permet pas d'utiliser ORCHIDEE couplé avec l'atmosphère, il faudrait

revoir le calcul du coefficient de traînée par exemple, ou encore la gestion de plusieurs bilans d'énergie de la surface terrestre par LMDZ. En effet, le profil de température est calculé de manière implicite dans le sol d'ORCHIDEE standard et dans LMDZ de manière couplée. Seul le bilan d'énergie est simulé, il n'y a pas de connexion avec le cycle de l'eau et donc avec le routage d'ORCHIDEE.

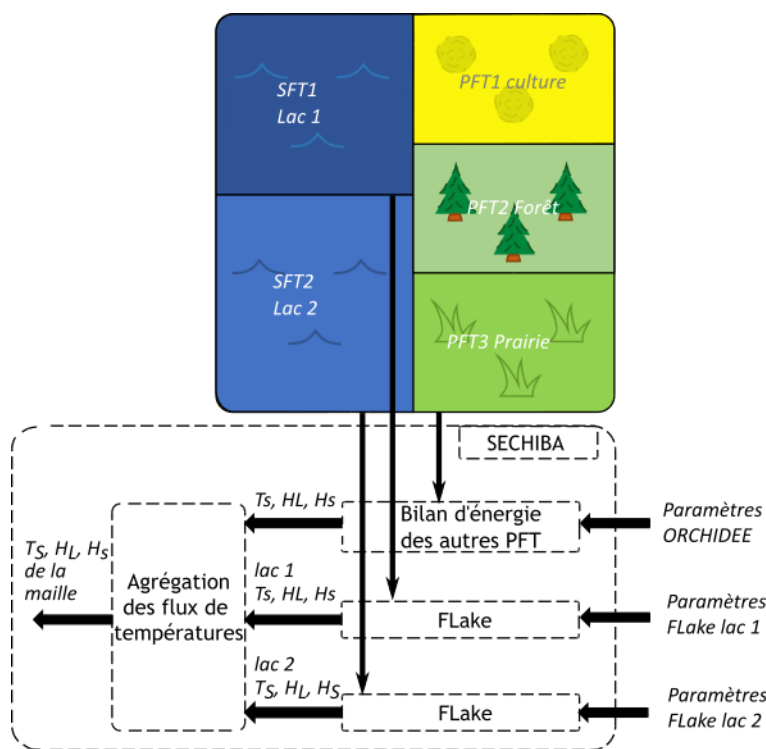


FIGURE 2.1 – Schéma de l'implémentation des lacs dans ORCHIDEE

## 2.4 . Présentation de FLake

### 2.4.1 . Concept de base

FLake est un modèle unidimensionnel qui simule le bilan d'énergie d'un lac. Il calcule le profil vertical de température en se basant sur la structure thermique naturelle des lacs. Cette structure est composée de deux couches qui sont la couche de mélange et la thermocline. Le profil de température est considéré uniforme dans la couche de mélange liée à la présence d'une turbulence élevée. La thermocline est stratifiée et suit un profil polynomial. Si les conditions extérieures l'exigent, le modèle peut faire geler le lac en modélisant une couche de glace et une couche de neige. Enfin, une couche de sédiment peut être activée pour les lacs peu profonds, permettant un échange de chaleur avec les sédiments.

Pour fonctionner, FLake a besoin des conditions météorologiques à la surface du lac (température, humidité relative, pression, vitesse du vent, rayonnement so-

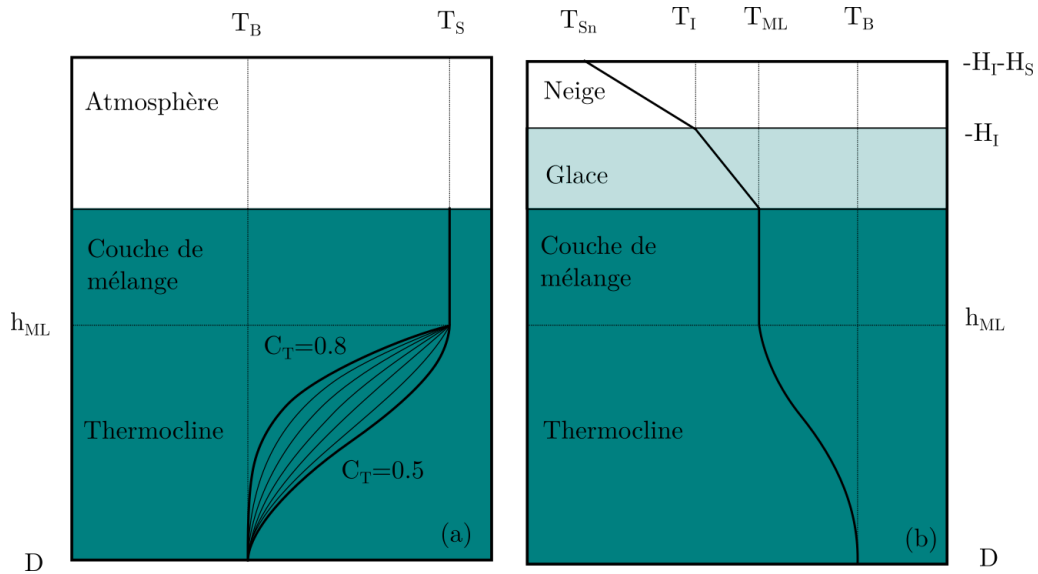


FIGURE 2.2 – Profil de température simulé par FLake (a) en eau libre et (b) pour un lac gelé (adapté de Mironov (2008))

laire et terrestre) et de paramètres physiques pour caractériser le lac. Le lac est représenté par sa profondeur moyenne, par un coefficient d'extinction qui permet de mesurer l'absorption de la lumière, l'albédo et le fetch.

#### 2.4.2 . Le profil de température dans la thermocline

La description du profil de température de chaque couche repose sur le concept d'auto similarité. Ainsi, le profil de température normalisée suit une fonction universelle  $\Phi_{\Theta}$  qui dépend de la profondeur normalisée  $\zeta$  (voir équation 2.11 et figure 2.2).

$$\Phi_{\Theta} = \frac{T_s - T}{T_s - T_b} \quad (2.11)$$

$$\zeta = \frac{z}{D}$$

où  $T_s$  la température de surface du lac,  $T_b$  est la température à l'interface entre les sédiments et le lac et  $T$  est la température à une certaine profondeur  $z$ .

L'idée d'une telle représentation émerge avec les travaux de Kitaigorodskii and Miropolsky (1970). Filyushkin and Miropolsky (1981) remarque que cette fonction dépend de l'état de la couche de mélange, plus précisément, si cette couche s'agrandit ou diminue. Les différentes paramétrisations de cette fonction universelle reposent sur le même principe que la méthode de Paulhosen dans la théorie de la couche limite. À savoir, un polynôme de degré 4 est choisi et les conditions aux limites permettent d'identifier les coefficients polynomiaux. La fonction universelle utilisée dans FLake provient de Malkki and Tamsalu (1985). Les deux premières

conditions imposent  $\Phi_{\Theta}(0) = 0$  et  $\Phi_{\Theta}(1) = 1$  pour retrouver  $T(0) = T_S$  et  $T(D) = T_b$ . La troisième condition est d'imposer la dérivée nulle de  $\Phi'_{\Theta}(1) = 0$ . Contrairement à d'autres travaux comme ceux de [Kitaigorodskii and Miropolsky \(1970\)](#) qui imposent aussi  $\Phi'''_{\Theta}(0) = 0$  et [Arsenev and Felzenbaum \(1977\)](#) qui ne retiennent que ces trois contraintes, la dernière condition est étalonnée à l'aide de mesures.

Finalement, la relation 2.12 est le polynôme qui représente le profil de température dans FLake.

$$\Phi_T(\zeta) = \begin{cases} 1 - (1 - \zeta)^3 & \text{si } dh_{ML}/dt > 0 \\ 1 - 4(1 - \zeta)^3 + 3(1 - \zeta)^4 & \text{si } dh_{ML}/dt \leq 0 \end{cases} \quad (2.12)$$

Dès que l'évolution de la couche de mélange change de direction, le passage d'un profil à l'autre n'est pas instantané mais lissé. Le modèle FLake utilise plutôt l'intégrale entre 0 et 1 du profil de température normalisé qui est le coefficient de forme  $C_T$  que le profil lui-même. À partir du coefficient  $C_T$ , on peut retrouver le profil de température avec le polynôme de l'équation 2.13. Il a été construit pour retrouver les équations 2.12 en y injectant les valeurs de  $C_T$  égales à 0.6 et 0.7.

$$\Phi_{\theta}(\zeta) = \left( \frac{40}{3}C_T - \frac{20}{3} \right) \zeta + (18 - 30C_T) \zeta^2 + (20C_T - 12) \zeta^3 + \left( \frac{5}{3} - \frac{10}{3}C_T \right) \zeta^4 \quad (2.13)$$

L'intervalle des valeurs prises par  $C_T$  a été élargi entre 0.5 et 0.8 pour augmenter l'intervalle de variation de  $\Phi_{\Theta}$ . Les différents profils de température peuvent être visualisés sur la figure 2.2. Pour faire évoluer le profil entre les 2 états, une loi simple (Équation 2.14) est établie sur la dérivée temporelle de  $C_T$ . C'est une constante qui dépend de l'intervalle de  $C_T$  et d'un temps caractéristique  $t_{rc}$ . Le signe de cette équation dépend du profil qui doit être atteint selon le sens de l'évolution de la couche de mélange.

$$\frac{dC_T}{dt} = \text{sign}(dh_{ML}/dt) \frac{C_T^{\max} - C_T^{\min}}{t_{rc}}, \quad C_T^{\min} \leq C_T \leq C_T^{\max} \quad (2.14)$$

### 2.4.3 . Les différents bilans d'énergie dans le lac

Pour définir, le profil de température de manière complète, il faut simuler la température de la couche de mélange et celle à l'interface avec les sédiments. Pour cela, un premier bilan d'énergie est établi sur la totalité du lac (équation 2.15). On note la profondeur moyenne du lac  $D$ , la température moyenne du lac  $\bar{T}$ , le flux de chaleur à l'interface entre le lac et les sédiments  $Q_b$ . Pour rappel,  $\rho_w$  est la masse volumique de l'eau (constante dans cette équation),  $c_w$  la capacité calorifique de l'eau,  $I(D)$  est le rayonnement solaire qui atteint l'interface lac-sédiment,  $SW_{\downarrow}$  le rayonnement solaire et  $LW_{\downarrow}$  le rayonnement terrestre qui atteint la surface.



$$D \frac{d\bar{T}}{dt} = \frac{1}{\rho_w c_w} [LW_{\downarrow} + (1 - \alpha_w)SW_{\downarrow} - Q_b - I(D)] \quad (2.15)$$

En reprenant la définition de  $C_T$  et en calculant la température moyenne  $\bar{T}$  du lac, on trouve l'équation 2.16.

$$\bar{T} = T_s - C_T(1 - h_{ML}/D)(T_s - T_b) \quad (2.16)$$

Un deuxième bilan d'énergie est nécessaire pour avoir une deuxième équation pour calculer  $T_s$  et  $T_b$ . On intègre donc l'équation de la chaleur sur la couche de mélange et on obtient l'équation 2.17. On note  $Q_h$  le flux de chaleur à l'interface entre la couche de mélange et la thermocline et  $h_{ML}$  la hauteur de la couche de mélange.

$$h_{ML} \frac{dT_s}{dt} = \frac{1}{\rho_w c_w} [LW_{\downarrow} + (1 - \alpha_w)SW_{\downarrow} - Q_h - I(h_{ML})] \quad (2.17)$$

Pour pouvoir simuler correctement le bilan d'énergie, il faut encore fermer les termes suivants : la hauteur de la couche de mélange  $h_{ML}$ , le flux entre la couche de mélange et la thermocline  $Q_h$ , le flux à la base de la thermocline  $Q_b$  et les flux de chaleur latente et sensible.  $Q_b$  est normalement calculé à l'aide du module de la couche de sédiment en considérant un autre bilan d'énergie dans cette couche avec un autre profil type dans cette couche. Si le module est désactivé,  $Q_b$  est considéré nul. La fermeture du terme  $h_{ML}$  sera explicité au paragraphe suivant. Les flux de chaleur seront aussi considérés dans un autre paragraphe de cette section. Il reste donc trois variables et il n'y a que les deux équations de bilan d'énergie. Il faut donc une dernière équation pour résoudre le problème.

On étudie les deux cas suivants : la couche de mélange se rétracte ou s'agrandit. Dans le premier cas, d'après les travaux de [Filyushkin and Miropolsky \(1981\)](#), le flux de chaleur turbulent  $Q$  dans la thermocline suit aussi le concept de similarité. Comme pour le profil de température  $\Phi_{\Theta}$ , on définit un profil de flux de chaleur turbulente  $\Phi_Q$  (Équation 2.18).

$$\Phi_Q(\zeta) = \frac{Q_h - Q}{Q_h - Q_b} \quad (2.18)$$

Cette nouvelle formulation du flux de chaleur est introduite dans le bilan d'énergie du lac. Elle est intégrée entre  $h_{ML}$  et une profondeur quelconque  $z'$  de la thermocline et puis intégrée une seconde fois sur la totalité de la thermocline par rapport à  $z'$ . On obtient alors l'équation 2.19 où  $C_{TT}$  est un nouveau coefficient de forme lié à cette double intégration et est égal à  $\int_0^1 d\zeta \int_0^{\zeta'} \Phi_T(\zeta') d\zeta'$  et  $C_Q$  est le coefficient de forme du profil de flux de chaleur qui est égal  $\int_0^1 \Phi_Q(\zeta) d\zeta$ .

$$\begin{aligned} & \frac{(D-h)^2}{2} \frac{dT_s}{dt} - \frac{d}{dt} [C_{TT}(D-h)^2(T_s - T_b)] \\ & = \frac{1}{\rho_w c_w} \left[ C_Q(D-h)(Q_h - Q_b) + (D-h)I(h) - \int_h^D I(z)dz \right] \end{aligned} \quad (2.19)$$

Une analyse physique suggère que  $C_Q = 2C_{TT}/C_T$ , l'annexe A du rapport de FLake explique cette relation.  $C_{TT}$  est facilement calculable en connaissant  $\Phi_\Theta$ .

Dans le deuxième cas où la couche de mélange se rétracte, la forme supposée du flux de chaleur turbulent n'est plus valable. Il faut donc trouver une autre équation. FLake considère alors la dérivée de  $T_b$  nulle.

Pour compléter les bilans d'énergie ci-dessus, il faut calculer les flux de chaleur sensible et latente. Pour cela, FLake calcule les flux moléculaires, de convections libres et turbulents. Le flux de chaleur latent et sensible est alors choisi parmi ces 3 en prenant le flux le plus élevé. Les flux turbulents sont calculés sur la base de la théorie de Monin-Obukhov présenté plus haut. Les formules de Charnock empiriques sont appliquées pour calculer la rugosité pour obtenir les flux turbulents en dissociant deux régimes de vitesse pour le vent : élevés et faibles.

#### 2.4.4 . Calcul de la hauteur de la couche de mélange

Deux cas sont considérés, soit la turbulence dans la couche de mélange est liée aux gradients de densité soit elle est induite par le vent. Dans le premier cas, le flux de flottabilité est négatif et l'évolution de la hauteur de la couche de mélange est régie par l'équation 2.20. Cette équation repose sur la même description que celle proposée par [Deardorff \(1970\)](#) pour de la couche limite atmosphérique instable. L'équation 2.20 décrit l'évolution de la couche de mélange en fonction des variables suivantes : une vitesse caractéristique  $w^*$ , la hauteur  $h_{ML}$  et un flux turbulent proche de la surface  $Q_s$ .

$$A + \frac{C_{c2}}{w^*} \frac{dh_{ML}}{dt} = C_{c1} \quad (2.20)$$

$A$  est le rapport entre le flux à la base de la couche de mélange  $Q_h$  avec un flux caractéristique  $Q^*$ . Ce rapport mesure l'efficacité de l'entraînement lié à la convection.  $C_{c1}$  et  $C_{c2}$  sont des constantes égales à 0.17 à et 1.0. À la différence de l'atmosphère, l'énergie absorbée dans la couche de mélange n'est pas négligeable par rapport au flux d'énergie turbulent aux frontières de celle-ci (i.e à l'interface air-eau et couche de mélange-thermocline). Ainsi le flux turbulent caractéristique tient compte de cette absorption (Équation 2.21). Finalement, le set de variables caractéristiques est  $h_{ML}$ ,  $Q^*$  donné par l'équation 2.21 et  $w^*$  donné par l'équation 2.22.

$$Q^* = Q_s + I_s + I(h_{ML}) - \frac{2}{h_{ML}} \int_0^{h_{ML}} I(z)dz \quad (2.21)$$

$$w^* = [-h_{ML}\beta (T_{Sn}) Q^* / (\rho_w c_w)]^{1/3} \quad (2.22)$$

$w^*$  est le même que dans la théorie de Deardoff à la différence que  $Q_s$  est remplacée par  $Q^*$ .

Dans le cas où le flux de flottabilité est positif et donc que le rayonnement solaire tend à stratifier la colonne d'eau, l'équation 2.23 permet d'obtenir la hauteur de la couche de mélange à l'équilibre.

$$\left(\frac{f h_{ML_e}}{C_n u^*}\right)^2 + \frac{h_{ML_e}}{C_s L} + \frac{N h_{ML_e}}{C_i u^*} = 1 \quad (2.23)$$

Cette équation est dérivée d'une analyse de la conservation du Energie cinétique turbulente (Turbulence kinetic energy) (TKE) en prenant en compte l'effet de la force de Coriolis, de la stabilité de la couche de mélange et du flux de flottabilité. Cette hauteur de la couche de mélange n'est pas atteinte instantanément. La relation 2.24 permet de quantifier l'évolution de celle-ci où  $h_{ML_e}$  est la hauteur de la couche de mélange à atteindre qui correspond à l'état d'équilibre et où  $t_{rh}$  est le temps caractéristique du processus.

$$\frac{dh_{ML}}{dt} = \frac{h_{ML_e} - h_{ML}}{C_{relax}} \quad (2.24)$$

#### 2.4.5 . Les couches de neige et de glace

Les couches de glace et neige sont aussi décrites par un profil de température normalisé et paramétrisé. Le profil de température est considéré linéaire dans la couche de neige et son profil suit l'équation 2.25.

$$T(z, t) = \begin{cases} T_f - [T_f - T_I(t)] \Phi_I(\zeta_I) & \text{pour } -H_I(t) \leq z \leq 0 \\ T_I(t) - [T_I(t) - T_S(t)] \Phi_S(\zeta_S) & \text{pour } -[H_I(t) + H_S(t)] \leq z \leq -H_I(t) \end{cases} \quad (2.25)$$

Les profils normalisés de la glace et de la neige sont notés respectivement  $\Phi_I$  et  $\Phi_{Sn}$  et les températures de gel et à l'interface neige-glace,  $T_f$  et  $T_I$ . Enfin, on note les hauteurs de la neige  $H_{Sn}$  et de la glace  $H_I$ . Le bilan d'énergie sur la couche de neige et de glace (Équation 2.26) et l'égalité des flux (Équation 2.27) à l'interface neige-glace permet de calculer les températures de surface (ici au sommet de la neige si elle est présente) et à l'interface neige-glace. Dans ces équations, on note  $\rho_i$  et  $\rho_s$  les masses volumiques de la glace et de la neige,  $c_i$  et  $c_s$  les capacités calorifiques massiques de la glace et de la neige et enfin  $\lambda_I$  et  $\lambda_{sn}$  les conductivités thermiques de la glace et de la neige. Pour rappel,  $I(z)$  est le rayonnement solaire qui atteint la profondeur  $z$  dans le lac.

$$\begin{aligned} \frac{d}{dt} \{ \rho_i c_i H_I [T_f - C_I (T_f - T_I)] + \rho_s c_s H_S [T_I - C_S (T_I - T_{Sn})] \} \\ - \rho_s c_s T_{Sn} \frac{d}{dt} (H_I + H_S) = Q_s + I_s - I(0) + \lambda_i \frac{T_f - T_I}{H_I} \Phi'_I(0) \end{aligned} \quad (2.26)$$

$$- \lambda_i \frac{T_f - T_I}{H_I} \Phi'_I(1) = -\lambda_s \frac{T_I - T_{Sn}}{H_S} \Phi'_S(0) \quad (2.27)$$

Il reste à calculer les hauteurs de glace  $H_I$  et de neige  $H_{Sn}$ . La neige dans le modèle FLake ne peut que fondre ou être accumulée via le forçage météorologique. La glace peut fondre ou provenir du gel de l'eau. En reprenant le bilan séparément de chaque couche (Équations 2.28 et 2.29), on peut obtenir la dérivée des hauteurs de neige et de glace qui vont permettre de prédire leur évolution. La masse volumique de la neige est paramétrée dans FLake à l'aide de la hauteur de la couche de neige, ainsi un terme supplémentaire apparaît dans le bilan de masse. Dans ces équations, on note  $L_f$  la chaleur latente de fusion de l'eau,  $Q_w$  le flux de chaleur à l'interface eau et glace,  $\left(\frac{dM_S}{dt}\right)_a$  correspond à l'accumulation neigeuse liée à la précipitation.

$$L_f \frac{d\rho_i H_I}{dt} = Q_w + \lambda_i \frac{T_f - T_I}{H_I} \Phi'_I(0) \quad (2.28)$$

$$L_f \frac{d\rho_s H_S}{dt} = - (Q_s + I_s) + I(-H_I) + L_f \left(\frac{dM_S}{dt}\right)_a + c_s T_f H_S \frac{d\rho_s}{dt} \quad (2.29)$$

Dans les différents bilans d'énergie ci-dessus, le rayonnement solaire incident est multiplié par l'albédo. Celui-ci est aussi paramétrisé en fonction de la température de surface à l'aide de l'équation 2.30. Cette équation demeure la même si c'est de la neige ou de la glace en surface. Seule la paramétrisation de l'albédo minimal  $\alpha_{min}$  et maximal  $\alpha_{max}$  peut changer. Par défaut, pour la glace (respectivement pour la neige),  $\alpha_{min}$  vaut 0.1 (resp. 0.1) et  $\alpha_{max}$  vaut 0.6 (resp. 0.6).  $C_\alpha$  est une constante empirique qui vaut 95.6.

$$\alpha = \alpha_{max} - (\alpha_{max} - \alpha_{min}) \exp[-C_\alpha (T_f - T_I) / T_f] \quad (2.30)$$

Dans les équations plus haut, la conductivité thermique et la masse volumique de la neige ne sont pas constantes, mais paramétrisées avec les équations 2.31 et 2.32 pour prendre en compte les changements liés à l'âge de la neige. Dans ces équations, on considère l'âge directement lié à la hauteur. Cette hypothèse est assez peu réaliste. On définit  $\Gamma_{\lambda_s}$  et  $\Gamma_{\rho_s}$  deux constantes empiriques égales à  $1.3 \text{ J} \cdot \text{m}^{-2} \text{ s}^{-1} \text{ K}^{-1}$  et à  $200 \text{ kg} \cdot \text{m}^{-4}$ .

$$\lambda_s = \min \{ \lambda_s^{\max}, \lambda_s^{\min} + H_S \Gamma_{\lambda_s} \rho_s / \rho_w \} \quad (2.31)$$

$$\rho_s = \min \left\{ \rho_s^{\max}, \frac{1 - H_S \Gamma_{\rho_s} / \rho_w}{\rho_s^{\min}} \right\} \quad (2.32)$$

## 2.4.6 . Améliorations apportées par rapport à la version standard de FLake

### 2.4.6.1 Paramétrisation de la couche de neige

Le module de neige de FLake présente quelques déficiences. Par exemple, [Mironov et al. \(2010\)](#) étalonne l'albédo de la couche de neige de la même manière que celle de la glace. Dans nos différents tests préliminaires, nous avons constaté la génération d'instabilités dans quelques cas. La température de la neige  $T_{Sn}$  est directement calculée à l'aide de l'équation d'égalité des flux à l'interface de la neige et de la glace comme évoqué dans le paragraphe précédent. Cette équation, appliquée au calcul de la température de la neige, est donnée par l'équation 2.33.

$$T_{Sn} = T_I + R_h \cdot (T_I - T_f) \quad (2.33)$$

$R_h$  est le rapport entre la résistance thermique de la couche de glace et de neige. En reprenant l'équation 2.27,  $R_h$  est égale à  $\frac{H_S \Phi'_f(1) \lambda_i}{H_I \Phi'_s(0) \lambda_s}$ . Les instabilités ont plus de chance de se produire avec une grande valeur de  $R_h$ . Ce qui en pratique correspond à une faible couche de neige.

[Semmler et al. \(2012\)](#) a proposé des améliorations à la paramétrisation de la température  $T_{Sn}$  et les a validé sur le lac de l'Ours au Canada. Les trois améliorations que nous avons implémentées sont les suivantes :

- Remplacer la paramétrisation de la densité de la neige par une valeur constante égale à  $320 \text{ kg/m}^3$
- Modifier la conductivité thermique de la neige avec l'équation 2.34
- Remplacer les valeurs maximales et minimales de l'albédo de la neige et de la glace

$$\lambda_{seff} = \lambda_s + (\lambda_i - \lambda_s) \cdot e^{h_s \cdot c_{eff}} \quad (2.34)$$

$\lambda_{seff}$  est la conductivité efficace de la neige qui prend en compte la propriété de la glace quand la hauteur de la couche de neige est faible.  $c_{eff}$  est une constante égale à  $5 \text{ m}^{-1}$ . La nouvelle paramétrisation de la conductivité de la neige tend à améliorer la stabilité du modèle. Les valeurs de  $\alpha_{min}$  deviennent pour la glace (respectivement pour la neige) 0.3 (resp. 0.77) et de  $\alpha_{max}$  0.5 (resp. 0.87). [Pietikäinen et al. \(2018\)](#) a utilisé cette paramétrisation pour coupler FLake avec le modèle régional REMO. Il a aussi utilisé un modèle plus sophistiqué pour calculer l'albédo de la neige que la simple paramétrisation de l'équation 2.30 de FLake. Pour supprimer les quelques instabilités restantes, l'accumulation de neige peut commencer si la couche de glace est assez grande (c'est-à-dire  $3 \text{ cm}$ ). Dans notre cas, nous avons fait le choix de réduire le pas de temps d'un facteur 50 pour simuler le bilan d'énergie dans les couches de glace et de neige quand nous utilisons le modèle

avec un pas de temps de 30 minutes. Ce qui a aussi pour effet de supprimer les instabilités même avec une couche de neige inférieure à 3 cm.

#### 2.4.6.2 Le coefficient de forme

Quelques problèmes ont été soulevés avec la paramétrisation de l'évolution du coefficient de forme  $C_T$ . Par exemple, [Salgado and Le Moigne \(2010\)](#) a mis en évidence qu'après quelques années de simulation, la température  $T_b$  reste figée sur une valeur et ce qui est expliqué par le fait que  $C_T$  reste à sa valeur maximale. Cela est lié à la présence d'oscillations non-physiques de  $C_T$  en hiver quand la couche de mélange atteint le fond du lac. Avec des simulations globales, on peut aussi voir cette anomalie dans certains lacs qui ne gèlent pas. [Salgado and Le Moigne \(2010\)](#) proposent de limiter l'évolution temporelle du coefficient  $C_T$  à 0.01 par heure dans l'équation 2.14 et d'élever la valeur minimale de  $C_T^{min}$  à 0.65 au lieu de 0.5. L'évolution de la température profonde est devenue plus réaliste après ces changements. Nous appliquerons aussi ces changements dans le couplage ORCHIDEE-FLake.

Une dernière piste d'amélioration de la thermocline qui n'est pas appliquée ici est de la paramétrer différemment pour prendre en compte une plus grande variété de profils de température. Cette idée est décrite par [Golosov et al. \(2018\)](#). Il reprend un polynôme de degré 4 pour décrire la thermocline (Équation 2.35).  $A(t)$  et  $B(t)$  sont similaire au nombre de Nusselt au sommet et à la base de la thermocline et sont donc égaux à  $A(t) = \frac{Q_h \cdot (D-h)}{\lambda_{h,eff} \cdot (T_S - T_D)}$ ,  $B(t) = \frac{Q_D \cdot (D-h)}{\lambda_{D,eff} \cdot (T_S - T_D)}$  où  $\lambda_{h,eff}$  et  $\lambda_{D,eff}$  sont les conductivités thermiques effectives de part et d'autre de la thermocline. Ils proposent donc de quantifier A et B selon des cas limites clairement identifiables. Dans un deuxième temps, ils proposent d'appliquer la même méthodologie sur ces deux paramètres que sur le coefficient de forme du modèle FLake pour faire évoluer progressivement le profil de température d'un état extrême à l'autre. Cette paramétrisation pourrait améliorer la représentation des lacs gelés par exemple.

$$\Phi_{\Theta}(\zeta) = A(t)\zeta + (3 - 2A(t) - B(t))\zeta^2 + (A(t) + B(t) - 2)\zeta^3 \quad (2.35)$$



## 3 - Analyse de sensibilité du modèle FLake

### 3.1 . Introduction

Nous présentons dans ce chapitre une étude qui a été publiée dans le journal JGR-atmosphère. Celle-ci décrit l'analyse de sensibilité du modèle FLake que nous avons menée dans l'optique de l'implémenter dans ORCHIDEE. Avec cette information, nous pouvons définir quels sont les paramètres importants qui sont influents sur les variables importantes du bilan d'énergie, à savoir la température de surface et les flux de température du lac. Ainsi, nous savons sur quels paramètres il faut se focaliser pour diminuer les erreurs liées à la paramétrisation.

### 3.2 . Méthode

Il existe principalement deux familles d'analyse de sensibilité. On recense d'abord les méthodes locales qui permettent de se focaliser sur la sensibilité du modèle autour de valeurs précises des paramètres d'entrée avec des variations très faibles. Ces techniques permettent par exemple d'étudier des effets de seuil. Les méthodes globales, quant à elles, étudient des intervalles plus larges des paramètres d'entrée.

Dans notre cas, nous avons utilisé la méthode de Sobol qui se base sur la décomposition de la variance. [Iooss and Lemaitre \(2015\)](#) décrivent les différentes techniques selon leur coût informatique et la complexité du modèle. La méthode de Sobol est un bon choix pour un modèle complexe et peu coûteux en temps de calcul et pour un set restreint de paramètres (au maximum une vingtaine).

Plusieurs méthodes se basent sur la décomposition de la variance, on les nomme de type ANOVA. On réalise de nombreuses simulations dans l'espace des paramètres d'entrée. Nous pouvons calculer la variance engendrée par toutes les simulations. Cependant, les méthodes ANOVA reposent sur le principe de décomposition où chaque terme de la variance se réfère à l'effet d'un paramètre ou de plusieurs paramètres (voir équation 2 de l'article). La particularité de la méthode de Sobol réside dans la façon d'échantillonner l'espace de départ avec un échantillonnage qui remplit de façon bien uniforme l'espace des paramètres. Finalement, les indices utilisés sont de la forme de l'équation 5 de l'article où les  $S_i$  sont les indices de Sobol qui correspondent au  $i^{\text{ème}}$  paramètre qui sont compris entre 0 et 1. Un indice proche de 1 correspond à un paramètre influent et proche de 0 très peu influent. Enfin, on définit les indices de Sobol totaux qui mesurent l'influence de tous les ordres de l'indice, c'est-à-dire à l'ordre 1, l'influence du paramètre, l'ordre 2 l'influence de celui-ci combiné à un autre paramètre, etc. Dans notre cas, nous avons des séries temporelles, nous avons donc aussi utilisé des indices de Sobol modifiés qui prennent en compte l'aspect temporel. En effet, la variance totale évolue avec le temps et il faut donner un poids plus faible aux périodes moins sensibles aux



paramètres. Donc ces indices sont écrits comme sous la forme de l'équation 7 de l'article.

Les différents paramètres étudiés dans l'analyse de sensibilité sont l'albédo de l'eau, le coefficient d'extinction de l'eau, la profondeur moyenne du lac et le fetch. Nous avons aussi choisi de rajouter un paramètre interne du modèle qui est  $t_{rc}$  qui est un temps caractéristique qui correspond au temps de relaxation du profil de température. Nous avons aussi rajouté l'albédo de la neige et de la glace ainsi que leur coefficient d'extinction respectifs. Nous avons désactivé la paramétrisation de l'albédo en la remplaçant par une valeur constante. En sachant qu'il existe des bases de données de la profondeur, nous avons décidé de garder les incertitudes liées à l'évolution temporelle des niveaux d'eau des lacs. Nous avons décidé de définir trois catégories de lacs selon leur profondeur en se basant sur la distribution globale des lacs.

Nous avons étudié 3 types de climat (tropical, tempéré et boréal) et 3 intervalles de profondeurs différentes. Pour l'intervalle des faibles profondeurs nous avons aussi étudié le cas avec la couche de sédiment activée. Nous avons utilisé le forçage WFDEI sur la période 2000-2004 avec un spin-up de 10 ans. Pour chaque cas, nous avons réalisé 11 000 simulations.

### 3.3 . Synthèse des résultats

Nos résultats confirment que la profondeur est le paramètre le plus influent sur la température de surface et les flux de surface. Cependant, la sensibilité de chaque paramètre évolue au cours des saisons. Pendant la saison de gel, l'albédo de la neige est le plus influent et lui seul contribue au premier ordre. Quand le lac n'est pas gelé, la sensibilité change selon l'intervalle des profondeurs considéré. Pour les lacs peu profonds, les incertitudes liées à la profondeur et au coefficient d'extinction expliquent la grande majorité de la variance de la température de surface qui est autour de  $2\text{ }^{\circ}\text{C}$  et du flux de chaleur latente qui est légèrement au-dessus de  $20\text{ W}/\text{m}^2$  dans les périodes chaudes quel que soit le climat. Pour les lacs les plus profonds, ce sont les incertitudes liées au coefficient de relaxation et à l'albédo qui expliquent la variance qui reste inférieure à  $1\text{ }^{\circ}\text{C}$  pour les températures et à  $10\text{ W}/\text{m}^2$  pour les flux de chaleur latente.

Cette analyse de sensibilité permet de montrer le comportement différent du modèle entre les lacs peu et très profonds. Ainsi, nous préconisons de séparer le traitement de ces lacs de manière différente dans les modèles de climat.

### 3.4 . Article 1

## Variance Based Sensitivity Analysis of FLake Lake Model for Global Land Surface Modeling

A. Bernus<sup>1</sup> , C. Ottlé<sup>1</sup> , and N. Raoult<sup>1</sup> 

<sup>1</sup>Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA-CNRS-UVSQ, Gif-sur-Yvette, France

### Key Points:

- FLake model parameters sensitivities on lake surface temperature and sensible and latent heat fluxes are analyzed
- Parameter sensitivities to surface variables vary with depth, climate, and season
- Respective roles of depth, fetch, surface albedo, light extinction, and thermocline shape coefficients are quantified

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

A. Bernus,  
[anthony.bernus@lscce.ipsl.fr](mailto:anthony.bernus@lscce.ipsl.fr)

### Citation:

Bernus, A., Ottlé, C., & Raoult, N. (2021). Variance based sensitivity analysis of FLake lake model for global land surface modeling. *Journal of Geophysical Research: Atmospheres*, 126, e2019JD031928. <https://doi.org/10.1029/2019JD031928>

Received 31 OCT 2019

Accepted 13 NOV 2020

### Author Contributions:

**Conceptualization:** C. Ottlé  
**Data curation:** A. Bernus  
**Formal analysis:** A. Bernus, C. Ottlé  
**Funding acquisition:** C. Ottlé  
**Investigation:** A. Bernus  
**Methodology:** A. Bernus, C. Ottlé  
**Project Administration:** C. Ottlé  
**Resources:** C. Ottlé  
**Software:** A. Bernus  
**Supervision:** C. Ottlé  
**Validation:** A. Bernus  
**Visualization:** A. Bernus, N. Raoult  
**Writing – original draft:** A. Bernus, C. Ottlé  
**Writing – review & editing:** A. Bernus, C. Ottlé, N. Raoult

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**Abstract** Given the ever increasing spatial resolution of climate models and the significant role of lakes on the regional climate, it becomes important to represent water bodies in climate models. Such developments have started in the IPSL (Institut Pierre Simon Laplace) climate model and its land surface component, ORganizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE), with the Freshwater Lake model, FLake. To answer the questions raised by these new developments, such as the lake differentiation and related model parameters, we analyze spatial distributions of lake characteristics in the whole world to perform a global sensitivity analysis of the FLake parameters. As a result, three different climates and four lake depth configurations were selected as test cases. The Sobol method as sensitivity analysis based on variance decomposition was chosen to rank parameters impact on the model output, that is, lake surface water temperature, latent and sensible heat fluxes. We focus on the 11 parameters of the FLake model, which are the lake depth, the albedo and light extinction coefficient of water, snow, and ice respectively, the fetch, and the relaxation coefficient of the thermocline shape factor. The results show different sensitivity features according to the lake type and climate. The dominant role and time varying contribution of the lake depth, radiative parameters (albedo, light extinction coefficient) and thermocline relaxation coefficient linked to the atmospheric conditions, were clearly highlighted. These findings will lead us to distinguish between different lake categories in each grid cell of ORCHIDEE in the future implementation.

**Plain Language Summary** Lakes are often neglected in climate modeling because their spatial extent does not exceed 4% of the Earth land area. But given the ever increasing spatial resolution of climate models and the significant role of lakes on the regional climate, it becomes important to represent water bodies in climate models. Such developments have started in the French IPSL climate model with the implementation of a one-dimensional freshwater lake model (FLake). To answer the questions raised by these new developments, such as the lake differentiation and related model parameters, we analyze spatial distributions of lake characteristics in the whole world to perform a global sensitivity analysis of the model parameters. Our results show the respective roles of the depth, fetch, thermocline, sediment, and radiative (albedo and light extinction coefficient) parameters in different climate conditions and at different time scales. Some avenues for further implementation and model parameters calibration are finally given.

## 1. Introduction

Lakes are often neglected in climate modeling because their spatial extent at global scale does not exceed 3.7% of the Earth's non glacial land area (Verpoorter et al., 2014). However, larger concentrations can be found in some regions. For example, Telmer and Costa (2007) studying two regions of 200 by 200 km in Canada and in Brazil, found that lakes cover respectively 33% and 19% of the land surface. In such conditions, lakes may have real impact on the atmospheric exchanges and consequently on weather prediction and climate modeling. Actually, the specific properties and physics of lakes (i.e., high heat capacity, low albedo and surface roughness, and water-ice transition phases) explain their major role on local meteorology and boundary layer stratification, as well as their influence on cloud formation (Laird et al., 2009). This has been demonstrated in numerous studies at global scale (Bonan, 1995; Dutra et al., 2010; Krinner, 2003; Mironov et al., 2010) and more specifically on the climate of regions where large fractions of land are covered by lakes. For example, Chuang and Sousounis (2003) showed the impact of the American/Canadian great lakes on the regional climate. They especially explained how processes associated with

these lakes deepen surface pressure, enhance cyclones, and consequently generate or strengthen a warm front. Song et al. (2004) and Thiery et al. (2015) studied the eastern African lakes and their influence. They notably showed the cooler effect of lakes during the day, the warmer one during the night, and how lakes strengthen precipitation. Even though the impacts on the energy and water cycles are more significant at regional scales, the importance of lakes on climate prediction at larger scales has also been shown by various authors (e.g., Bonan, 1995 or Le Moigne et al., 2016). As an example, using the CNRM-CM5 climate model, Le Moigne et al. (2016) showed the substantial role of lakes on the simulation of regional surface energy budgets and the reduction of some air temperature biases mainly over the region of the Canadian Great Lakes.

The role of lakes and especially of thermokarst lakes on the methane global cycle and its climate feedbacks is also an important question. Bastviken et al. (2004) pointed out the important lake contribution to the non-anthropogenic methane emissions and consequently on climate. They estimated that between 8% and 16% of non-anthropogenic methane emissions come from lakes, and West et al. (2016) specified that the larger contribution arises from shallow and eutrophic lakes.

Given these features and the continuous increased spatial resolution of general circulation models, many efforts have been pursued during the last decade to include lake models in numerical weather prediction (NWP) and climate models. A good trade-off between numerical cost and the correct representation of the surface main processes is offered by one dimensional (1-D) approaches. Between the crude representation of bulk models representing lakes as a single water column completely mixed (Ljungemyr et al., 1996) and multi-layer approaches resolving the transport equations for the turbulence kinetic energy and its dissipation rate (Stepanenko, 2005), conceptual models such as the FLake model (Mironov, 2008) appear to be a good compromise able to simulate the lake water surface temperature (LWST) and the thermocline profile with a two-layer resolution of the energy budget.

FLake is a well-tested model and its ability to predict the temperature structure of various lakes has been demonstrated in numerous studies (Voros et al., 2010; Woolway & Merchant, 2019). FLake has been implemented in numerous climate and NWP models. To cite but a few, FLake is part of the German COSMO and TERRA land component (Mironov et al., 2010), the European ECMWF Earth System Model and its HTESSEL land surface component (Balsamo et al., 2012; Dutra et al., 2010), the UK Met Office Unified Model and the JULES Land simulator (Rooney & Bornemann, 2013), and the French CNRM-CM5 and its SURFEX interface (Le Moigne et al., 2016; Salgado & Le Moigne, 2010). All these works have demonstrated the added value of the FLake model in better simulating regional weather and climates.

When implemented at large scale, a lake model requires the definition of a lake fraction in each grid cell of the land surface model. Generally, the grid cell size is a few tens of kilometers for a climate model. Such information can be offered by existing global databases generally derived from remote sensing data. These databases provide water bodies at a very fine spatial resolution, according to the decametric resolution of optical and microwave sensors. As an example, the product of Pekel et al. (2016), derived from Landsat imagery, provides the water extent of each water body with a resolution of 30 m, during 32 years 1984–2015 and at a monthly time scale.

In addition to determining the different types of water surfaces that need to be considered, the definition of the model parameters necessary to initialize and run the model is an important task. In this aim, the prior model sensitivity analysis (SA) is a powerful tool that has to be conducted on the basis of previous knowledge about the spatial/temporal variabilities of lake features. Previous studies based on FLake model have shown the dominant role of the lake depth (Balsamo, 2010) and of the extinction coefficient of solar radiation (Heiskanen et al., 2015). The lake bathymetry has definitely a strong impact on the temporal dynamics of its temperature and capacity to mix and freeze. This property has been used by Balsamo (2010) to derive successfully, lake depth by optimization of LWST with observations. More recently, Layden et al. (2016) showed that not only lake depth but also snow and ice albedos and light extinction coefficient could be optimized in the same way globally, with space thermal infrared measurements. Although these studies have demonstrated the dominant role of lake depth and radiative properties on LWST, in other words their sensitivities, to our knowledge, no quantitative SA of FLake output variables has been conducted until now.

The work presented in this study aims to fill this gap. It is part of a community project devoted to the inclusion of water bodies in the IPSL climate model and more precisely in its land surface component: the ORCHIDEE land surface model. Previous studies to represent lakes in ORCHIDEE were focused either on the atmospheric impacts with no connections with the other hydrological processes (Krinner, 2003; S. Peng, personal communication, 2015) or on their impacts on the water flows via the routing hydrological scheme (d'Orgeval et al., 2008; X. Zhou, personal communication, 2019). In our future developments, we aim to represent the lakes' water and energy budgets explicitly by the addition of a lake tile in the ORCHIDEE Multi-Energy Budget version under development at IPSL. For this purpose, the FLake model has been chosen to resolve the surface energy budget on the water fraction of each grid mesh. However, before running the model, a number of parameters need to be prescribed and calibrated. To help this implementation in ORCHIDEE, we performed a preliminary SA of the FLake model in standalone mode. SA is an effective tool used to identify the dominant processes and understand their dynamics. It allows to rank the parameters and to determine which ones should require the largest attention in the following calibration step (Saltelli, 2008).

In this study, we therefore show the results of a global sensitivity analysis of the FLake model designed with this aim of its forthcoming global scale implementation. Using three representative climates, we therefore focus on three of its main output variables at the atmosphere interface, that is, LWST and the turbulent latent ( $H_L$ ) and sensible ( $H_S$ ) heat fluxes. The goal of our study is then to answer the following questions:

- What are the main characteristics of lakes at a global scale?
- How do we define the study cases and the parameter space for this sensitivity analysis? Given the parameter space, which parameters are the most influential?
- How should model implementation be done to account for this diversity of lake properties in parameterization?

## 2. Materials and Methods

### 2.1. FLake Model

FLake (Mironov, 2008) is a 1-D thermodynamic lake model developed for NWP 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, and wind speed). 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 sediment layer allows for the representation of the heat transfer at the lake bottom-sediment interface, which is not negligible in the case of shallow lakes. The structure of the lake thermocline is parameterized 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 (see for examples Figures 3 and 4 in Mironov, 2008 for a representation of the depth-temperature relationships).

The resolution of the bulk energy budgets of the mixed layer and thermocline allows to calculate the model prognostic variables, that is, the mixed-layer temperature and depth, the free water bottom temperature, the thermocline shape factor, the ice layer top temperature, and ice thickness. The resolution of the energy budget of the mixed layer allows to estimate the latent and sensible heat fluxes at the water-atmosphere interface.

When a sediment layer is accounted, two more variables are predicted: the attenuation depth of the annual thermal wave and the temperature at that depth. The bottom temperature and the total depth of the sediment layer are prescribed. If the snow module is activated, the snow temperature at the atmosphere interface and the snow layer thickness are calculated, based on the bulk resolution of the heat budget of the snow layer. It should be noted that the model resolves only the energy budget equations and that the water balance is not solved. That means that the water volume is kept constant in time, the lake depth and surface extent are therefore input parameters of the model. In addition, two important radiative parameters (surface albedo and light extinction coefficient both for water, snow

and ice) and 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 with a constant temperature near of 4°C. To run the model, we add a false bottom by shortening the lake depth of deep lakes to 50 m. This allows us to get around the modelization of the hypolimnion. Also, the absence of 2-D and 3-D processes (e.g., lake currents) has shown the limits of FLake's performances in certain cases. For example, FLake fails to capture the typical pattern of springtime warming in the deep North American Great Lakes (Martynov et al., 2010). However, such examples remain specific and the overall performance of FLake is promising, especially given its low numerical cost.

For the implementation of FLake in the ORCHIDEE land surface model, a multi-tiled version of the model is used, for which separate energy budgets can be calculated for each type of surface and especially for water bodies. In order to prepare the lake model developments and the necessary step of parameter calibration, a global SA has been designed and performed.

## 2.2. Sobol Method

SA studies the variation of the model simulated variables according to the variations of some of its parameters. Here, we perform a global SA where each parameter varies over the entire range of variation. As such, we get information for the whole parameter space compared to local SA techniques where we would only get information around one point. Also, the parameters in a global SA can vary simultaneously, whereas this is not the case for local SA experiments where the parameters vary one by one (Saltelli, 2008).

First, we need to define the model inputs, parameters, forcing, and outputs not to mix up these concepts. The inputs are the external data needed to run the model. These can take the form of parameters, which describe explicitly the system (here the lake). Alternatively, forcing variables are inputs that provide the boundary conditions of the model (here the interface between the lake and the atmosphere). The atmospheric forcing inputs are the meteorological variables used to drive the model. These variables are not taken into account during the SA. Finally, the outputs are the variables that the model calculates.

To quantify the model sensitivity, we used the Sobol method. The main principle is to decompose the variance of the considered outputs into variances explained by the variation of each input. During this process, the parameter sampling should be independent from each other.

Briefly, the method is decomposed into four steps. The first one is to list the parameters and define their range of variation. Each parameter will be associated with a random variable uniformly distributed on each interval. The second is to consider the parameter as random variable and to sample the intervals (with Monte Carlo method for example) in a way to allow the variance decomposition estimation. The third step is then to perform the simulations needed and to choose a scalar output; this can either be the model output directly if it is already a scalar or a value calculated from the model output to create a scalar, for example, the Root Mean Square Error between modeled time series and observations. The final step consists in the analysis of the model output variance according to the parameter ranges.

This method is widely used in geophysical sciences and especially by the land surface model community. For example, Xu et al. (2012) applied this method to study the parameter sensitivity of the Common Land Surface Model (CLM). Li et al. (2013) also performed a sensitivity analysis on CLM for six different output variables using several SA methods including the Sobol one. In the same way, Zheng et al. (2019) studied the sensitivity of evapotranspiration and runoff to the Noah-MP land surface model parameters, and Guerrero et al. (2017) studied the sensitivity of the surface fluxes calculated by the Canadian Small Lake Model to its input parameters, in the specific case of a small lake in Canada. To understand how the model output variance can be decomposed with the Sobol approach, we first introduce some definitions. If we define a random variable (or stochastic variable) for each parameter with a uniform probability distribution on each interval, we can represent a model under a functional form:

$$Y = f(\mathbf{X}) = f(X_1, \dots, X_p) \quad (1)$$

where  $Y$  is the stochastic variable corresponding to the model output that we want to study and  $\mathbf{X} = (X_1, \dots, X_p)$  is the parameter set with  $p$  parameters.  $\mathbf{X}$  is a random vector with a uniform distribution on their interval definition. As mentioned previously,  $Y$  is a scalar and so  $f$  is either the model output or the composition between the model and a post-processing metric.

We define  $V(Y)$ , the variance of the random variable  $Y$ ,  $E(Y)$ , the expected value of  $Y$ , and  $E(Y|X_i)$ , the condition expected value of  $Y$  given  $X_i$ .

Sobol' demonstrated that the function  $f$  can be decomposed as in Equation 2. All the terms in this equation are orthogonal (namely the integral of the multiplication of two different terms is null), and the integrals of  $f_i$  over  $X_i$  interval,  $f_{ij}$  over  $X_i$  and  $X_j$  interval, etc, are null. This decomposition exists and is unique if all the random variables are independent.

$$f(\mathbf{X}) = f_0 + \sum_{i=2}^p f_i(X_i) + \sum_{i=1}^p \sum_{j>i}^p f_{ij}(X_i, X_j) + \dots + f_{12\dots p} \quad (2)$$

It can be shown that the terms of Equation 2 can be written using Equation 3 with  $I = \{i_1, i_2, \dots, i_s\}$  be a set of  $s$  subscripts and  $\hat{I}$  be the concatenation of these subscripts, that is,  $i_1 i_2 \dots i_s$ :

$$f_{\hat{I}}(X_{i_1}, \dots, X_{i_s}) = E(Y | X_{i_1}, \dots, X_{i_s}) - \sum_{\{Z \in \mathcal{P}(I) | Z \neq I\}} f_{\hat{Z}} \quad (3)$$

where  $\mathcal{P}(I)$  denotes the power of set of  $I$ , that is, all subsets of  $I$ , including  $I$  itself and the empty set. By convention, if  $I$  or  $Z$  is  $\emptyset$ , then  $f_I$  or  $f_{\hat{Z}}$  is equal to  $f_0$ .

This formula is coherent with  $f_0$  being constant and equal to the expectation values of the output. The first term of Equation 3,  $E(Y | X_{i_1}, \dots, X_{i_s})$  is a conditional expected value, and the last term removes all inferior orders so that it only accounts for the interaction between the  $s$  parameters used.

From Equation 2, the total variance of function  $f$ ,  $V(Y)$ , is decomposed into component variances from individual parameters and their interactions.

Each variance component is the variance of the terms of Equation 2. If the variance of the sum of all the terms is the sum of the variance of each term, it is because each term is orthogonal with each other.

$$V(Y) = \sum_{i=1}^p V_i + \sum_{i=1}^p \sum_{j>i}^p V_{ij} + \dots + V_{12\dots p} \quad (4)$$

where  $V_{\hat{I}} = V(E(Y | X_{i_1}, \dots, X_{i_s})) - \sum_{\{Z \in \mathcal{P}(I) | Z \neq I\}} V_{\hat{Z}}$ .

From the previous equation, each term can be divided by the total variance to create the so-called Sobol indices. Each term measures the sensitivity of the output  $Y$  for each parameter  $X_i$  or the interactions between several and is written:

$$S_{\hat{I}} = \frac{V_{\hat{I}}}{V(Y)} \quad (5)$$

The first-order effect indices  $S_i$  denotes the sensitivity resulting from the main effect of individual parameter  $X_i$ . The second-order effect indice  $S_{ij}$  indicates the sensitivity resulting from the interaction of two parameters  $X_i$  and  $X_j$ . The higher order effect indices are not often calculated for their high numerical cost and their low utility. Another indice can be developed: the total effect sensitivity indice. It represents the main effect of  $X_i$  and all its interactions with the other parameters. This indice is defined as:

$$S_{T_i} = 1 - \frac{V_{1,\dots,i-1,i+1,\dots,p}(E(Y | X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p))}{V(Y)} \quad (6)$$

All the Sobol indices have values between 0 and 1, and the sum is equal to 1 whereas the sum of the total indices may be superior to 1 if the variance  $Y$  linked to parameter interactions is not null.

In our case, the variable  $Y$  evolves with time. A preliminary analysis could be to apply the Sobol method on the  $Y$  averaged over a time period, typically one year. However, the different mechanisms of the study model varying temporally, this analysis would not be sufficient. In order to explore parameter sensitivities at various time scales, Lamboni et al. (2011) proposed the generalized Sobol indices defined by:

$$GS_i = \frac{\sum_{t=1}^N V(Y(t))}{\sum_{t=1}^N V(Y(t))} S_i(t) \quad \text{and} \quad GS_{T_i} = \frac{\sum_{t=1}^N V(Y(t))}{\sum_{t=1}^N V(Y(t))} S_{T_i}(t) \quad (7)$$

where  $\sum_{t=1}^N V(Y(t))$  is the variance of  $Y$  over a time period of length  $N$  and  $V(Y(t))$  the variance of  $Y$  at the time step  $t$ .

These generalized Sobol indices are very interesting because they allow to show synthetic information at different time scales for example monthly or annual. In the case of a lake model, because of the water phase changes, density and thermal properties, we expect the radiative and convective processes to have time varying relative weights, and so, the influence of the parameters to change with time. Therefore, in this study, the Sobol indices have been also analyzed at different time scales from the annual to the hourly one.

To calculate the model output variances and to derived the Sobol indices, we follow the method presented by Saltelli et al. (2010). Then, we define two matrices  $\mathbf{A}$  and  $\mathbf{B}$  with  $p$  columns and  $N$  lines where  $N$  is the sample size. The component  $a_{ij}$  represents the value of the parameter  $i$  for the simulation  $j$ .

We also use the notation  $[\mathbf{B}]_{1 \leq k \leq p, j}$  for the line  $j$  of the matrix  $\mathbf{B}$ ,  $k$  is the column index. Each line corresponds to one model simulation. We also define  $\mathbf{A}_i$ , equal to matrix  $\mathbf{A}$  where the  $i$ th column has been replaced by the corresponding  $i$ th column of the  $\mathbf{B}$  matrix. To generate these matrices, we use the quasi Monte Carlo method. The sampling is generated by low discrepancy series and here we use the Sobol ones (Sobol, 1990). Then, the estimation of the Sobol indices is done with the following equations:

- For the first order:

$$S_i = \frac{1}{N} \sum_{j=1}^N f([\mathbf{B}]_{1 \leq k \leq p, j}) (f([\mathbf{A}_i]_{1 \leq k \leq p, j}) - f([\mathbf{A}]_{1 \leq k \leq p, j})) \quad (8)$$

- For the total order:

$$S_i = \frac{1}{2N} \sum_{j=1}^N (f([\mathbf{A}]_{1 \leq k \leq p, j}) - f([\mathbf{A}_i]_{1 \leq k \leq p, j}))^2 \quad (9)$$

Although Sobol method has intensive computational requirements, its sensitivity indices have been shown to be more effective than other approaches in capturing the interaction effects of several parameters for highly nonlinear models (Tang et al., 2007). Overall computing the first-order and total-order sensitivity indices requires:

$$N_{sim} = N \times (p + 2) \quad (10)$$

with  $N_{sim}$  is the number of simulations,  $N$  is the sample size, and  $p$  is the number of parameters.

### 2.3. Experimental Design

The experiments were driven by two items: our final objective to model all the different kinds of lakes in ORCHIDEE, from the boreal thermokarst shallow lakes to the large deep lakes of temperate and tropical regions, and the knowledge of the existing parameter database to estimate the parameter uncertainties.

In order to do so, we separately apply the Sobol method on LWST,  $H_L$ , and  $H_S$  variables, calculated by FLake over a set of 12 representative lakes ranging different depths and generic climates. Therefore,  $Y$  in the previous section is alternatively LWST,  $H_L$ , and  $H_S$ .

The Sobol method is applied separately for each test case, considering their specific parameter intervals and meteorological forcing. We choose several cases to reflect the diversity in climate and lake size found around the world. We therefore decided to first explore various global databases of lake depth and extent in

order to select the representative lake depths and extents of the global distribution on which to perform the SA. Then, for each representative case, we define the uncertainty of each parameter, based on the available datasets and known literature, to create the intervals and hence the parameter space used for the study.

### 2.3.1. Representative Study Cases

We used the HydroLAKES database (Messenger et al., 2016) to investigate the global distribution of natural lakes and human-made reservoirs in extent and lake depth. This database provides polygons of water bodies with surface areas of at least 10 ha and their main morphometric features including average lake depth and volume.

We analyzed the distribution of the average lake depth of each lake weighted by their surface extent (see Figure S1). In this figure, a mode around 3 m corresponding to shallow lakes (lakes with depth inferior to 5 m represent about 30% of the total lake area) can be observed, as well as an important contribution of the very deep lakes to this distribution (about 30% of which will be all simulated by FLake with the recommended prescribed depth of 50 m). These features have been already discussed in various works (Messenger et al., 2016; Pekel et al., 2016). Given these features, we decided to retain three typical lake depths in the following SA: 3 and 50 m, and we chose an intermediate case at 25 m to represent the case of medium lake size.

### 2.3.2. Climate and Meteorological Conditions

Lakes are present in various regions under diverse climatic conditions. We therefore decided to perform our SA on the main three climates (boreal, temperate, and tropical) of the Earth. In order to get realistic but generic meteorological forcing and to explore the sensitivities at different time steps, we use the WFDEI atmospheric reanalysis (Weedon et al., 2014) provided at a spatial resolution of 0.5° and 3-hourly time step, to force our model. We linearized the atmospheric forcing to get a time step of 30 min in order to model the diurnal cycle and assess the model sensitivities at hourly scale.

Three meteorological time series were extracted for a 5-year period (2000–2004) to represent the meteorological conditions of three regions of the world influenced by lakes: Central Siberia, Central Europe, and East Africa. These points were selected to get representative meteorological conditions over our representative lakes in these three main types of climate.

### 2.3.3. FLake Parameter Space

The first step in a SA is the definition of the parameter space, that is, the study parameters and their range of variation. In our case, it is based on the parameter uncertainties. As previously noted, we want to study the parameter sensitivity in three climates for three typical lake depths. Since the sediment layer has proved to be influential for shallow lakes, we propose to study 12 representative cases resulting from the combination of the three typical climates (boreal, temperate, and tropical) and the four lake types (shallow, shallow with sediments, intermediate, and deep).

The selection of model parameters was done based on previous studies, notably ones using LWST to calibrate the model parameters (Layden et al., 2016). We retain the usual input parameters of FLake (lake depth, extinction coefficient and albedo both for free water, ice and snow, fetch, sediment depth, and bottom temperature) and add the relaxation constant governing the time evolution of the thermocline shape factor, whose value is highly uncertain (Salgado & Le Moigne, 2010; Layden et al., 2016; D. Mironov and G. Kirillin, personal communication, 2019). Thus, the vector  $\mathbf{X}$  of the method section corresponds to a vector of random variables of nine-dimensions (when the sediment layer is not activated) and 11 dimensions in the other cases.

The ranges of variation that have been prescribed in the SA were defined according to the following approach. When FLake is run at global scale, these parameters are prescribed to default values, constant with time and space except for lake depth and fetch for which global maps of lake extent and lake depth are available. We therefore assumed that the uncertainties are mostly the result of the non-accounted spatial or temporal variabilities. Then, we defined the parameter ranges of variation from the analysis of existing databases providing such information, as explained in the following. The final ranges of variation are given in Table 1 for the three kinds of lakes (shallow, intermediate, and deep).



**Table 1**  
Parameter Ranges of Variation Used in the SA of the FLake Model

Parameter	Symbol	Shallow (3 m)	Intermediate (25 m)	Deep (50 m)
Lake depth (m)	$D$	[0.5, 5.5]	[22.5, 27.5]	[47.5, 52.5]
Fetch (m)	$F$	[0, 1300]	[0, 10600]	[0, 26,600]
Light extinction coefficient of water ( $\text{m}^{-1}$ )	$\kappa_w$		[0.225, 2.435]	
Light extinction coefficient of snow ( $\text{m}^{-1}$ )	$\kappa_s$		[15, 25]	
Light extinction coefficient of ice ( $\text{m}^{-1}$ )	$\kappa_i$		[8.4, 17.1]	
Water Albedo	$\alpha_w$		[0.025, 0.175]	
Snow Albedo	$\alpha_s$		[0.4, 0.8]	
Ice Albedo	$\alpha_i$		[0.4, 0.65]	
Relaxation coefficient (log)	$C_{\text{relax}}$		[2, 5]	
Sediment <sup>a</sup> :				
bottom temperature (K)	$T_H$	boreal	[271.2, 273.2]	N/A
		temperate	[282.7, 284.1]	N/A
		tropical	[295.0, 296.8]	N/A
depth (m)	$H$	[3, 10]		N/A

<sup>a</sup>Two different types of shallow lakes are considered: with a sediment layer and without. These parameters are used when the sediment layer is activated.

- Depth: We use the HYDROWEB database (Crétau et al., 2011) available at <https://www.theia-land.fr> to estimate the range of uncertainty of our three categories of lakes. We estimate them as the annual lake depth amplitude. HYDROWEB provides surface water levels (and not lake depth) from satellite altimetry on large lakes and some virtual stations on the biggest rivers on the Earth. Time series of lake water levels are provided for 160 lakes, and we have used these data to estimate the maximal annual amplitude of the water level variations. By subtracting the annual minimum level to the annual maximum level, we can assess the lake depth amplitude. The results show that the annual amplitude is quite variable between lakes, varying between 20 cm and 19.5 m (after filtering the inconsistent data), due to the variability of the water balance processes. Given that our perspective in the future is to model the mass balance, we decided to take as uncertainty range, the average annual amplitude, which is equal to 5 m

- Light extinction coefficients: for free water, the database ILEC (ILEC: World Lake Database, International Lake Environment Committee Foundation, 1999) has been chosen to estimate the uncertainty of the light extinction coefficient, following the work of Layden et al. (2016). ILEC provides *in situ* measurements of the Secchi depth on 150 lakes all over the world. The Secchi depth can be related to the lake turbidity with empirical formula. We used the equation proposed by Poole and Atkins (1929) to derive the extinction coefficient. This formulation, which relates linearly the Secchi depth to the inverse of the extinction coefficient, has been derived from numerous measurements. The analysis of the data distribution allows us to estimate a single range for our study cases equal to (0.23–2.44)  $\text{m}^{-1}$ , which corresponds to the first and third quartiles of the distribution. As a point of comparison, the values suggested in FLake vary from 0.2  $\text{m}^{-1}$  (for a very clear lake) to 2  $\text{m}^{-1}$  (very turbid case) (Mironov et al., 2010). For the ice, the range of variation of the light extinction coefficient was provided by Launiainen and Cheng (1998) following the FLake documentation. The blue ice is considered as the most transparent with a value of 8.4  $\text{m}^{-1}$  and the white one as the most opaque with a coefficient of 17.1  $\text{m}^{-1}$ . For snow, the range of variation was set to (15–25)  $\text{m}^{-1}$  depending on its liquid water content with the dry snow considered as the most opaque

- Albedos: For free water and in absence of global databases, we refer to literature values and essentially to McMahon and Moore (2017). The final interval has been set to (0.025–0.175), which can be compared to the default value of 0.07 used in FLake (the value commonly used for the ocean albedo). For snow and ice, since the default values proposed by Mironov et al. (2012) (i.e., a single and common interval equal to [0.1–0.6]) was quite different from the values extracted from literature review, we decided to follow Le Moigne et al. (2016) for snow and Ebert and Curry (1993) for the ice. Therefore, the final ranges of variations were set to (0.4–0.8) for snow and (0.4–0.65) for the ice albedo
- Fetch: The fetch length is the length of water over which a given wind has blown without encountering obstacles. Over lakes, it can be estimated from the lake area assuming a given shape. We use the HydroLAKES database (Messenger et al., 2016) to estimate the fetch value for each lake referenced. Assuming a circle shape, the fetch length is derived from the circle diameter. The fetch intervals are centered on the mean diameter for each kind of lake (shallow, intermediate, and deep). Then, we add the fetch error which has been derived from an ellipse construction. In this way, the fetch can be a value between the two ellipse axes, both determined from the surface area and shoreline length for each lake, depending on the wind direction. Finally, the error is approached by the mean difference of the two axes for each lake distribution (shallow, intermediate, and deep).
- Thermocline shape factor: As already noted by Layden et al. (2016) and Salgado and Le Moigne (2010), the relaxation coefficient, which is part of the prognostic equation of the thermocline shape factor, is very uncertain. The default value is  $3 \times 10^{-3}$ , but values from  $10^{-2}$  to  $10^{-5}$  were adjusted for small to deep lakes. We kept this range of variation in our SA following Layden et al. (2016).
- Sediment parameters: when the sediment layer is activated, two other parameters are required and therefore added in the SA. They are the sediment layer depth and bottom temperature. The range of variation of the depth has been set to the interval (3–10) m (see <http://www.flake.igb-berlin.de/usefulhints.shtml>), and the bottom temperature has been derived from the annual air temperature averaged over a 15-year period (1990–2004) for our three climates following the advice of Golosov and Kirillin (2010). The range of variations were derived from the interannual variability calculated over the same time period.

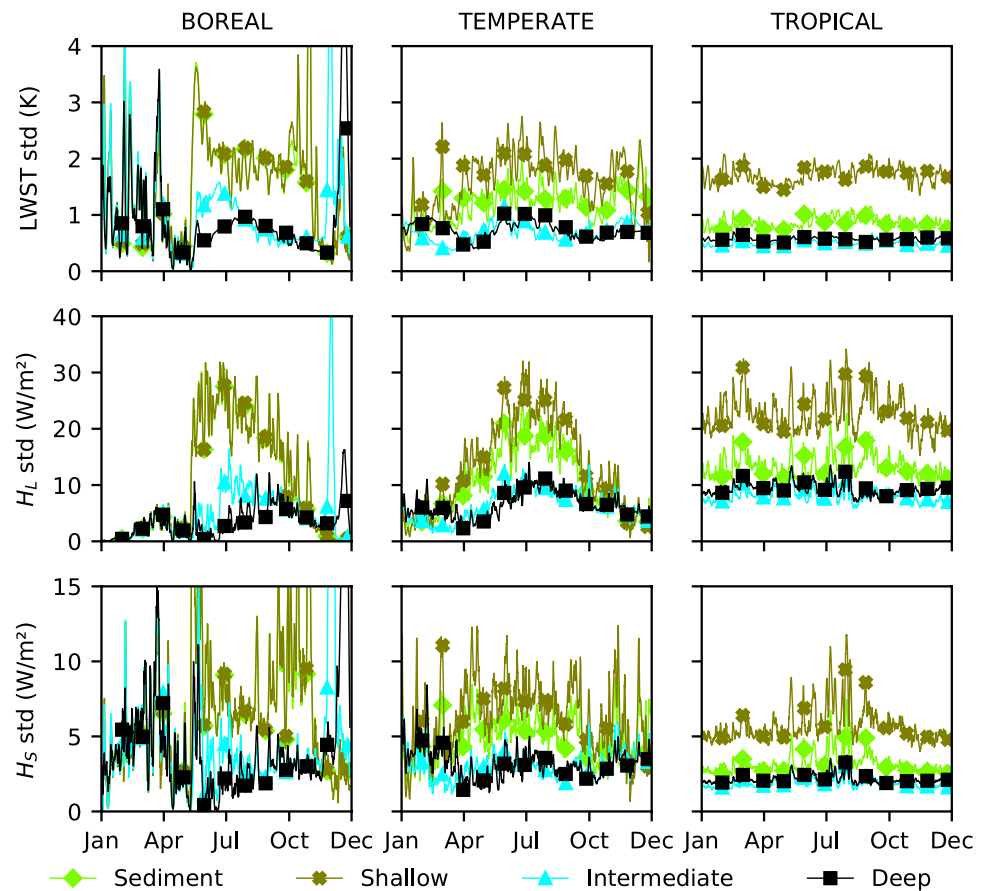
#### 2.3.4. Simulations Protocol

To generate the FLake simulations for the 12 study cases (three climates and four lake types), we have chosen to sample the parameter space with a uniform distribution using a quasi Monte Carlo algorithm based on low discrepancy Sobol series. The number of simulations have been defined following Equation 10 with  $N$  set to 1000. As a result, for the different lake cases, 11,000 (when 9 parameters are sampled) or 13,000 simulations (in the sediment cases with 11 parameters) were generated. In comparison, Rosolem et al. (2012) performed 45,000 simulations for 42 parameters and Zhang et al. (2013) generated 60,000 for 28 parameters.

Each simulation was generated on a fifteen-year period (1990–2004). Ten years of spin-up are necessary and sufficient to reach model equilibrium and get rid of incorrect initial conditions. Hence the SA has been performed on the last 5 years of the simulation (2000–2004). The variance analysis and Sobol decomposition have been performed on the three main output variables of the model: LWST,  $H_L$ , and  $H_S$ . Since the conclusions are generally the same for these three variables, we generally plot only the ones concerning LWST, the others are given in the supplementary document.

### 3. Results

The model parameter sensitivities are shown in the following for LWST,  $H_L$ , and  $H_S$ . The sensitivities were studied at annual, monthly and daily scales through the Generalized Sobol Indices (GSI) and the daily Sobol Indices (First order and Total order). We recall that all the Sobol indices are in the range (0, 1), that the sum among the nine or the 11 parameters (sediment case) is equal to 1 and the larger the value, the greater the sensitivity. Since these indices are relative to the total variance of the studied variable, it is important to look first to the simulated variances of our main surface variables of interest, that is, LWST,  $H_L$ , and  $H_S$ .



**Figure 1.** Total variances of LWST,  $H_L$ , and  $H_S$  simulated by our ensemble of simulations for the four studied lake types over the Year 2001. Each column corresponds to one of the three climate conditions. The time series are calculated at half-hourly time step and smoothed over a 1 day averaging window.

The experiments were conducted over a 5-year study period. However, for sake of clarity, and because the results do not differ from one year to the other, we only plot the results obtained for the year 2001 in this section.

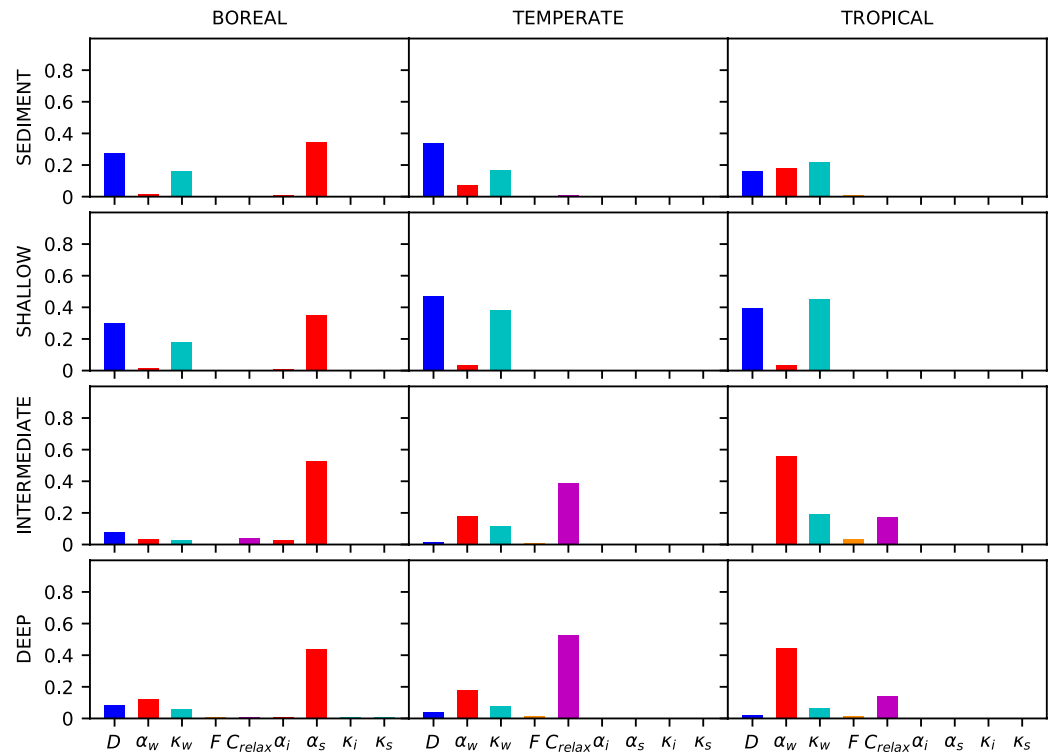
### 3.1. Total Variances of Lake Surface Variables

Figure 1 shows the total variances of LWST,  $H_L$ , and  $H_S$  generated by our ensemble of simulations for our 12 representative cases. For each case, the total variances are calculated at a time step of 30 min with a smoothing average of 1 day. The figure shows clearly that the shallower the lake, the larger the variances. Indeed, LWST variances for shallow lakes are more than twice those of deep lakes and can exceed 3 K in summer boreal conditions.

For climates with a strong seasonality, for example, boreal, the variances are larger in summer. In addition, for the intermediate/deep lakes in boreal conditions, large peaks are observed in fall and winter. These peaks can be explained by the larger variability of the frozen conditions among the ensemble of simulations.

The impact of the sediment layer on LWST is larger under tropical climates and during summer in temperate regions. The addition of the sediment layer results in the variances halving. This effect is similar to the effect of increasing the lake depth.

For the simulated variances of  $H_L$  and  $H_S$  (Figure 1), the same features can be seen. However, larger values are obtained for  $H_L$  compared to  $H_S$ . This also results in a more pronounced seasonal cycle for  $H_L$  in summer for the temperate and boreal climates.



**Figure 2.** The first order generalized Sobol indices are presented in vertical bar graphs. Columns (respectively lines) are related to climate conditions (respectively lake types). Each bar represents the parameter contribution to the LWST variance for Year 2001.

### 3.2. Generalized Sobol Indices

Figure 2 shows the GSI for each of the nine parameters shared by all the model configurations. The GSI are calculated for LWST at annual time scale. The relative importance of each parameter changes depending on lake type and climate.

For shallow lakes (with and without sediments), the dominant effect of lake depth and water light extinction coefficient on LWST is clear. The influence of the light extinction coefficient of water is more important for temperate and tropical climates compared to the boreal one. Under boreal conditions, the snow albedo parameter dominates because of the frozen and snowy conditions present during several months of the year. The other parameters are not influential for these shallow lakes, except for the water albedo which appears to be more sensitive when the sediment layer is activated under temperate and tropical climates.

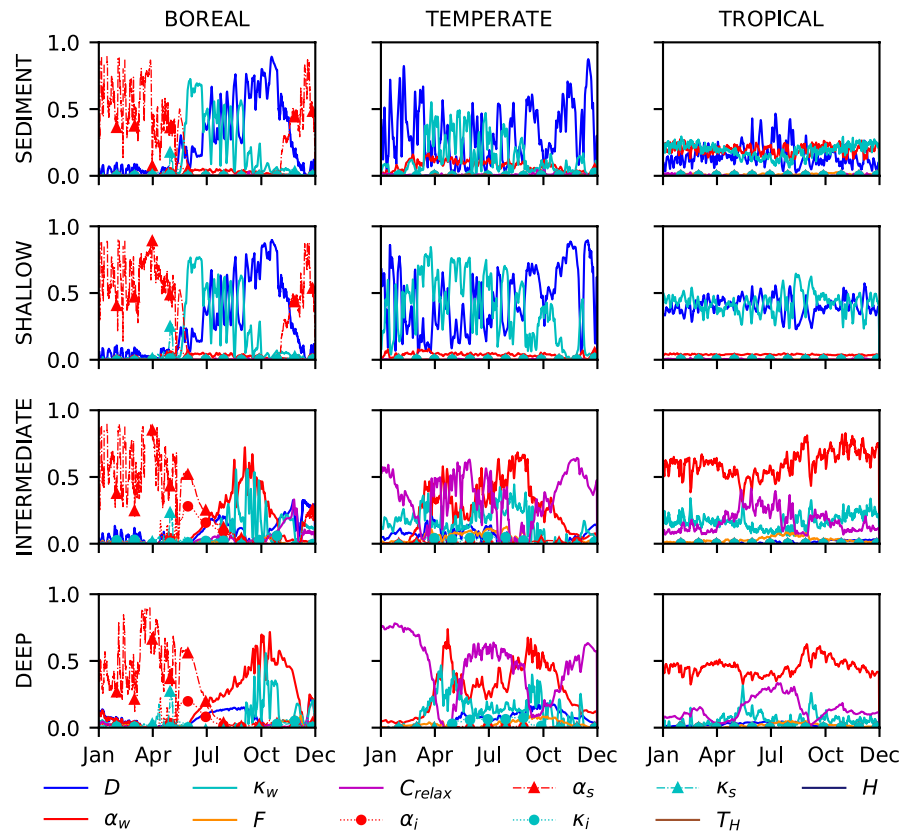
For intermediate/deep lakes, the parameter contributions are different. The relaxation time of the thermocline shape factor and water albedo (snow albedo for the boreal climate) dominates instead of the lake depth and water light extinction coefficient. Similarly, to shallow lakes, water albedo is most important in tropical climates where solar radiation is larger.

We further calculated the same indices on  $H_L$  and  $H_S$  (supplementary document: Figures S2 and S3). The same features are observed, with the exception of the fetch parameter that appears to be slightly sensitive in the case of deep lakes.

### 3.3. Sobol Indices at Daily Time Step

#### 3.3.1. First Order Indices on LWST

Figure 3 shows the temporal series of the first order SI (for LWST), calculated at a time step of 30 min and smoothed over a daily window for each of the 11 model parameters and for each of the 12 lake cases.



**Figure 3.** LWST first order Sobol indices are shown here for the 12 study cases and at a half-hourly time step (after application of a 1-day moving average). Columns (respectively lines) are related to climate conditions (respective lake types). Colors are related to model parameters, lines to free water, triangle markers to snow optical parameters and circles to the ice ones.

The noisy character of the curves is explained by the temporal variability of the meteorological conditions (cloudiness, wind speed, air temperature, and vapor pressure), which impact the total variance of LWST and the relative contribution of the lake physical processes. Despite these high frequency variations, the plots show the same main features already identified in the GSI plots (Figure 2); (i) the key role of the snow/ice radiative parameters in frozen conditions, (ii) the dominant influence of lake depth and water light extinction coefficient for shallow lakes, and (iii) the larger influence of fresh water albedo and relaxation coefficient for deeper lakes. The seasonal variations also bring additional information. An annual cycle is clearly highlighted, more important under boreal climates than tropical ones, because of the changing atmospheric conditions throughout the seasons.

For shallow lakes (both with and without sediment), the respective roles of lake depth and water light extinction coefficient appear to be complementary. During fall and winter (under unfrozen conditions), the lake depth has a larger role than water light extinction coefficient, whereas in summer, the water light extinction coefficient gains importance. In contrast, these two parameters have a more balanced influence under tropical conditions where there is no seasonal effect.

For intermediate/deep lakes, fresh water albedo and water light extinction coefficient have the same temporal evolution, which acts opposite to the evolution of the thermocline relaxation coefficient, especially under temperate and tropical climates. The temporal variation of the SI of the thermocline relaxation coefficient, which vanishes twice per year in spring and fall, is particularly remarkable. This is due to the time variation of the dominant physical processes that occur in the temperate and tropical lakes all along the year. Lakes show thermal stratification, which evolves seasonally according to the surface energy balance at the atmosphere interface because of the water density variation with temperature and

its anomaly at 4°C (larger density). The lake mixing resulting from the convection flows can occur at a number of times during the year depending mostly on lake depth and atmospheric conditions (such lakes are called holomictic). When the lake is mixed, the temperature is homogeneous, and the thermocline is non-existent. The shape factor will therefore be non-efficient and its impact null. On the contrary, when the lake is stratified (in summer or in winter for temperate lakes), the thermocline develops, and the shape factor plays a larger role.

As expected, the fresh water albedo plays a larger role when the solar radiation is larger, that is, for tropical climates and during summer for boreal and temperate climates. It seems also to play a larger role in shallow lakes when the sediment layer is activated, in temperate and tropical conditions. In fact, this feature highlights the lower role played by the water light extinction coefficient and the lake depth when the sediment processes are accounted for in shallow lakes.

The addition of a sediment layer with the energy transfers at the interface impacts the surface temperature because of the added constraint on the lake bottom temperature. As a result, the thermocline temperature is affected and the main parameters driving it (in the shallow lake case, water light extinction, and lake depth) see their influence decrease. The fresh water albedo which is the third ranked influential parameter on LWST, sees its sensitivity increasing.

In the case of frozen lakes, the ice and snow albedos are the dominant parameters because of their strong opacity, which do not allow the water below to play any role on the surface-atmosphere processes. The interplay of the ice/snow radiative parameters during the snow melting season in spring and for the intermediate/deep lakes is clearly seen on the plots and explained by the longer length of the ice melting period compared to that of shallow lakes. The lake depth is dominant during the fall season, since it greatly controls the lake capacity to freeze.

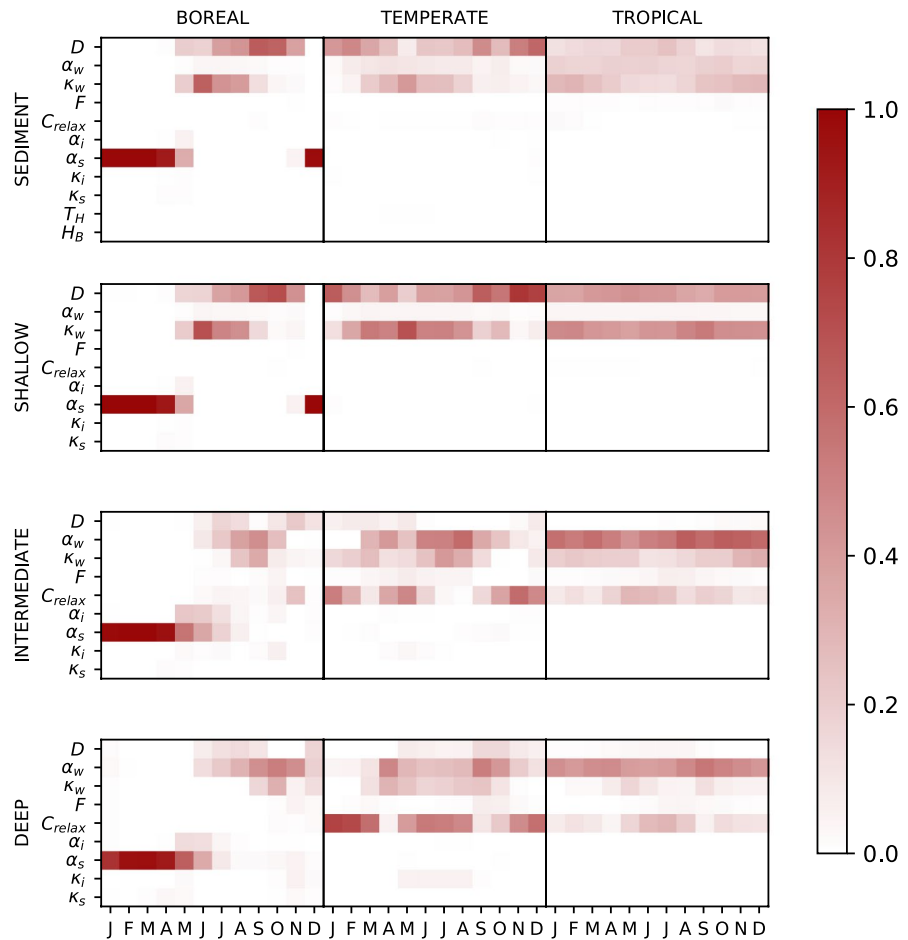
A synthetic picture of the parameter sensitivities all along the year is provided by the representation of the first order GSI calculated at monthly timescales (Figure 4). Indeed, the figure summarizes the features already observed at daily scale, that is, the dominant role of the snow/ice albedos in frozen conditions, of depth and light extinction coefficient for shallow lakes, and the larger contribution of water albedo and relaxation constant for deeper lakes. Their seasonality is clearly shown also, according to the seasonality of the dominant physical processes (see, e.g., the sensitivity drops of the relaxation constant during the turnover periods for the dimictic temperate lakes).

All these features suggest therefore some correlations between parameters, at least between lake depth and water light extinction coefficient in the case of shallow lakes that can be explored with the higher order Sobol indices.

### 3.3.2. Higher Order Indices on LWST

In Figure 5, the differences between the total order and first order SI are plotted for LWST. By subtracting the first order SI from the total order SI, we are left with the higher order SI, which quantify the interactions between the parameters. When two parameters dominate this difference, we can infer that the variance is influenced by the correlation between these parameters. As it can be seen on Figure 5, for shallow lakes, these differences are small (less than 0.2) and concerned only the lake depth and water light extinction coefficient. This means that a very small part of LWST variance is explained by the second order interactions between these two parameters. When the sediment layer is activated and in free water conditions, lake depth and water light extinction appear more correlated with second order effects explaining about half of the variance under temperate and tropical climates. The rest of the variance is explained by the higher order interactions between the sediment bottom temperature, the fetch, and in the case of tropical conditions, the relaxation constant (in the case of tropical conditions only). This could be the consequence of the thermal processes described previously and the impact of the sediment bottom temperature on the mixed layer temperature. The energy exchanges at the lake-sediment interface, driven partly by the imposed bottom sediment temperature, influence indirectly the weight of the other parameters influencing the depth of the mixed layer and of the thermocline like the fetch and the relaxation constant.

For deeper lakes, the correlations between the parameters become larger but are quite negligible in summer season. The differences between total and first order indices are much lower (generally less than 0.2), and

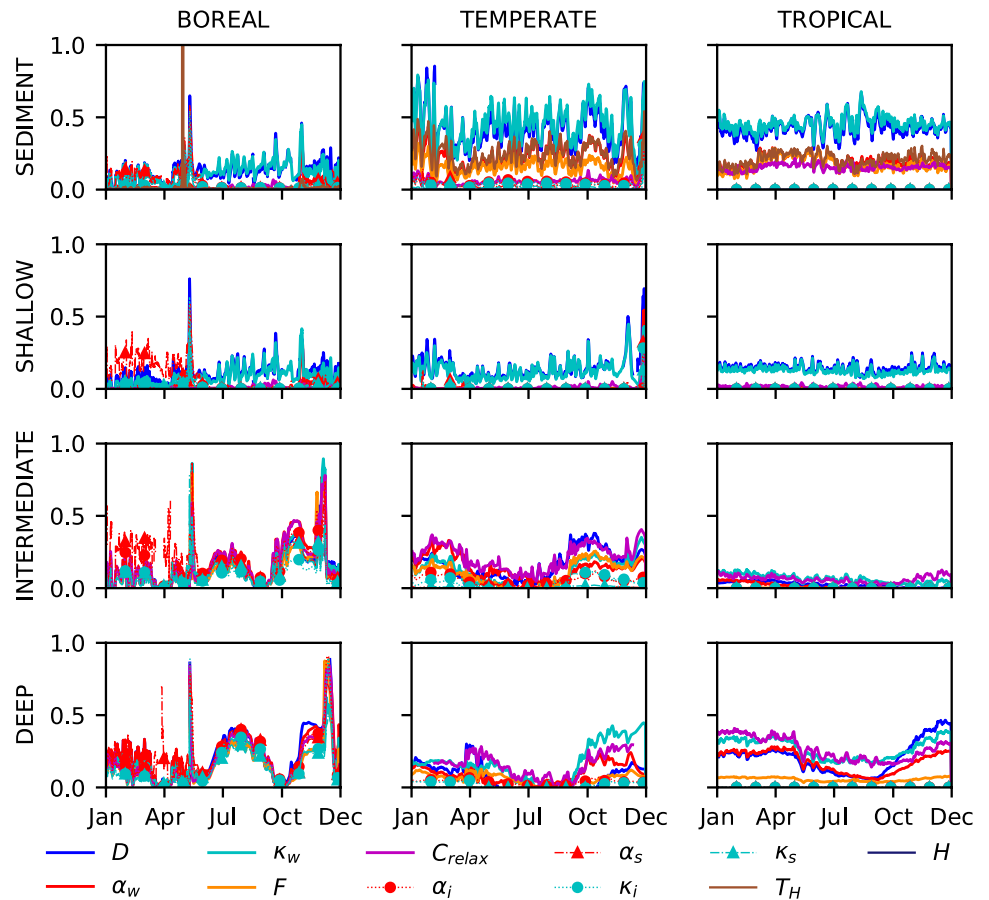


**Figure 4.** The first order GSI at monthly time step are shown for the 12 study cases. Each matrix column (respective line) corresponds to a month (respective parameter). Columns of layout (respective lines) are related to climate conditions (respective lake types). Color intensity is related to the parameter influence.

their behaviors suggest some correlations between water light extinction and relaxation constant on one side and between fresh water albedo and lake depth on the other side. For deep lakes, we also observe larger values in winter under boreal conditions. However, these values are simply due to the larger variability of the frozen conditions and reflect the combined impacts of lake depth, fresh water albedo, fetch and extinction coefficient on the lake stratification and frozen status.

### 3.3.3. Sensitivity on Surface Turbulent Fluxes

Figure 6 shows the first order Sobol indices calculated for  $H_L$  for our 12 study cases (the results for  $H_S$  are very similar and can be found in Figure S4). The same features already observed for LWST are visible here. Lake depth and water light extinction are the dominant parameters for shallow lakes in unfrozen periods, with the fresh water albedo playing a larger role in the case of the added sediment layer. In freezing conditions, the snow/ice albedos are the most influential as already noted for LWST. For deeper lakes, the fresh water albedo and shape factor relaxation coefficient are dominant. However, the fetch parameter appears more sensitive especially during the periods of lake turnover in spring and fall. This is explained by the dominant role of wind in the mixing processes during these periods. Concerning the parameter interactions, exactly the same features observed on LWST can be seen on the surface fluxes, as expected.



**Figure 5.** Differences between the LWST total and first order Sobol indices at a half hour time step (after application of a 1-day moving average) for the 12 study cases and for Year 2001. Columns (respective lines) are related to climate conditions (respective lake types). Colors are related to model parameters, lines to free water, triangle markers to snow optical parameters and circles to the ice ones.

## 4. Discussion

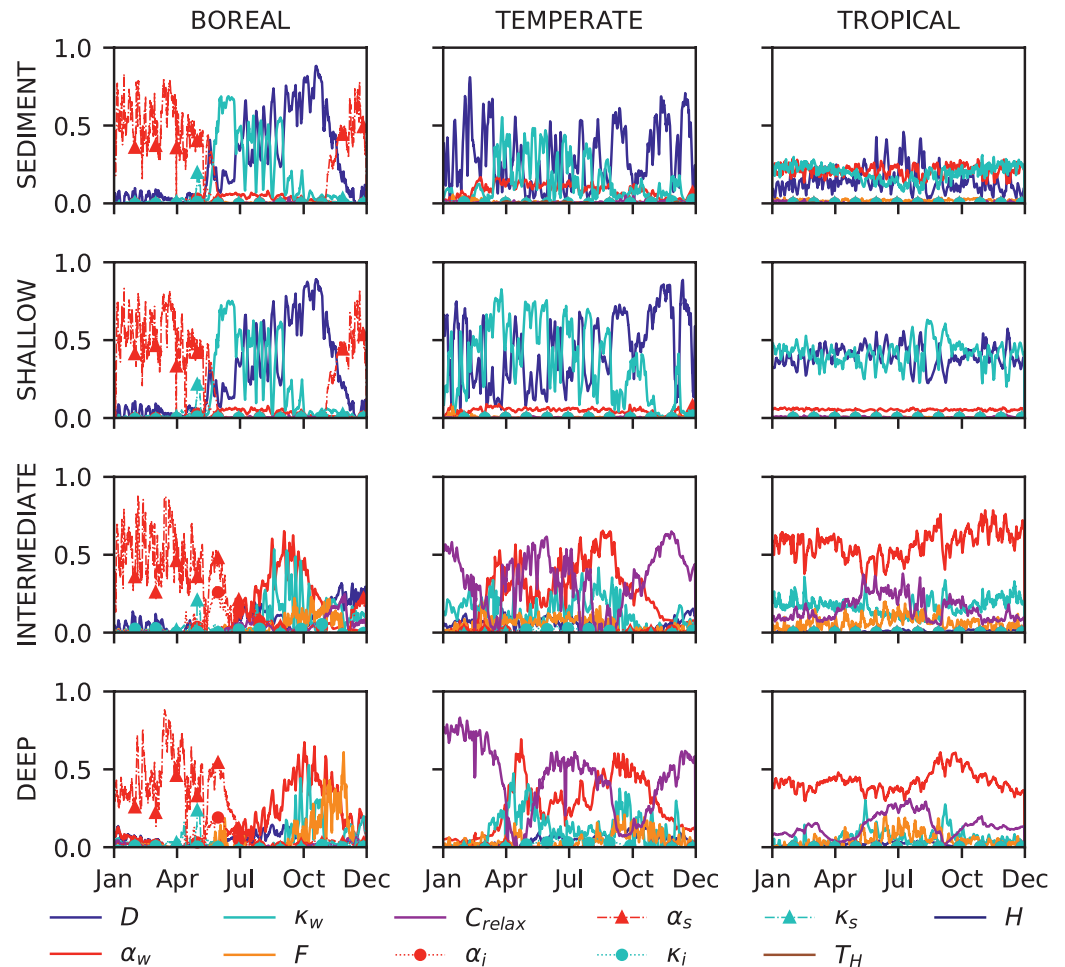
### 4.1. Respective Role of Model Parameters

First of all, our results confirmed the dominant influence of the lake depth and the varying influence of the other parameters on the thermal processes. Due to the seasonal and diurnal variations of the atmospheric conditions, the dominant parameters vary along the year. In frozen conditions, the snow and ice albedos were shown to play a key role on the surface temperature and fluxes as expected. Furthermore, the dominant parameters are not the same when shallow or deep lakes are considered.

For shallow lakes, in the case of free water, it has been shown that lake depth and water light extinction uncertainties explain the larger part of the variance of LWST and the surface fluxes. For intermediate/deep lakes, the total variance is more explained by the fresh water albedo and thermocline shape relaxation coefficient. The larger sensitivity of the surface fresh water albedo for deep lakes compared to shallow ones was not shown in previous studies. Although the final impact on LWST and fluxes is quite moderate (less than 1 K for LWST, 5 W/m<sup>2</sup> for  $H_S$ , and 15 W/m<sup>2</sup> for  $H_L$ , respectively), this parameter should receive some attention in the future when more measurements will be available.

The water light extinction coefficient is less influential on LWST and heat fluxes for deeper lakes. For these lakes, there is generally a deep mixed layer. In FLake, the water light extinction coefficient influences essentially the amount of solar energy that will be transferred in the thermocline below the mixing layer. If the mixing layer is deep, it will absorb all the solar energy no matter the value of the light extinction coefficient.





**Figure 6.**  $H_L$  first order Sobol indices are shown here at a half-hourly time step (after application of a 1-day smoothing average) for the 12 study cases. Columns (respective lines) are related to climate conditions (respective lake types). Colors are related to model parameters, lines to free water, triangle markers to snow optical parameters and circles to the ice ones.

These results are in line with the findings of Heiskanen et al. (2015). In their study, the same argument was given to explain why this coefficient is only influential when the lake is very clear and not turbid. This explains also the non-negligible correlations that we identified between the lake depth parameter and water light extinction coefficient, and the role of lake depth on the mixing layer thickness leads to secondary effects on the sensitivity of the water light extinction coefficient.

In this study, the model is not sensitive to the fetch, even for the deepest lakes where a larger parameter range is prescribed. Yet, the role of fetch is well-known on water mixing. This lack of sensitivity is because the effect of the fetch is not fully represented in FLake since it is a 1D model, and the large convection flows are not represented.

As with any sensitivity analysis experiment, these results are dependent on the definition of the parameter space, that is, the selected parameters and their range of variation. As an example, consider the water light extinction coefficient discussed above. If the minimum of the variation range was set to  $1 \text{ m}^{-1}$  instead of  $0.2 \text{ m}^{-1}$ , that is, excluding very clear cases, then, even in the case of shallow lakes, the SI values would be close to 0. In the same way, other internal parameters fixed constant could have been considered in this study, such as the relaxing constant for the wind mixed layer. Our strategy was to focus our analysis on the parameters that were already identified by FLake community as subject to uncertainties and

to define their range of variation according to the expected errors in a future implementation at global scale. As a consequence, our SA results are more relevant for the global application foreseen. Different applications of FLake would require dedicated SA focused on the uncertain parameters and output variables of interest.

In this work, we have also tried to assess the role of the sediment parameters (depth and bottom temperature) on LWST and surface fluxes. We showed that the influence of these two parameters is very negligible compared to the other model parameters at the first order. This is an advantage since these parameters are very uncertain. However, this result has to be taken carefully because the ranges of variation that we have prescribed may be underestimated, especially for the bottom temperature. In our work, and in absence of measurements, this temperature has been estimated from the annual air temperature following the advice of Golosov and Kirillin (2010). Such assumption, if not valid, could produce incorrect heating or cooling flux at the soil-sediment interface and lead to significant errors on the thermocline temperature and, to a lesser extent, on the surface variables.

#### 4.2. Global Modeling Perspectives

The results obtained through this SA allowed us to answer our preliminary modeling issues. First, the analysis of lakes global features showed the spatial distribution of lakes and their main features according to spatial extent, seasonal dynamics and most importantly lake depth. From the HydroLAKES database, we have seen that deep lakes represent more than 30% of the total lake area globally and that shallow lakes with lake depth lower than 5 m represent about 30% of the distribution (this number could be however underestimated because of the limited resolution of the sensors used to map water bodies. As an example, this limit is equal to 10 ha for HydroLAKES (Messenger et al., 2016). Their spatial distribution told us that the density of shallow lakes is clearly dominant in the northern latitudes, whereas the deep lakes can be distributed all over the world. Given these different features and parameter sensitivities, it appears important to distinguish these two kinds of lakes during our implementation of FLake into ORCHIDEE.

Second, the larger sensitivity of shallow lakes to lake depth confirms also the added contribution of a continuous monitoring and in this objective, the importance to simulate the lake water budget and the relations with surface extent. In this perspective, the benefit of future altimetry missions, such as the SWOT satellite based on radar interferometry, should be of great interest to initialize and update regularly our parameterizations. SWOT will be able to provide surface extent and water levels of all the water bodies of size larger than 6 ha, at global scale and with a repetitivity of at least 21 days (Biancamaria et al., 2016). New products of lake turbidity and albedo are also coming out in the remote sensing community, such as those derived from Sentinel-2 instruments (<https://www.theia-land.fr/en/products/>), and they should be helpful to better assess the light extinction coefficient as well as the albedo and their respective ranges of variation. Data assimilation of multi-datasets (albedo, turbidity, LWST, surface water elevation, and surface extent) is therefore a promising track to calibrate and update regularly lake model parameters.

Third, this study highlighted as well, the contribution of the relaxation constant of the thermocline shape factor on the surface temperature of deep lakes. However, the dynamics of this shape factor is hard to model. This issue was already raised by different authors such as Salgado and Le Moigne (2010) who proposed a solution to limit its high frequency time variations. Another solution was proposed by Golosov et al. (2018) but more work is needed for calibration. As a first step, we will follow Salgado and Le Moigne (2010)'s approach, which is easier to implement, in order to avoid some of the unrealistic behaviors they revealed during the lake turnover periods. Further work is currently being undertaken by G. Kirillin (personal communication, 2019) to better parameterize the relaxation constant according to lake depth.

Finally, the notable impact of the sediment processes on the surface variables despite the low sensitivity of the related parameters was shown in this study. Given the large uncertainties associated with these parameters in the absence of information on the sediment characteristics, we will not activate this module in FLake in its global scale implementation.

## 5. Conclusion

To answer some questions raised by the foreseen implementation of the FLake lake model into ORCHIDEE, a SA of model parameters was applied to four different kinds of lakes under three different climates. The analysis was performed on the 11 parameters of FLake that will need to be prescribed at global scale (lake depth, albedo and light extinction both for fresh water, ice and snow, fetch, thermocline shape relaxation coefficient, sediment layer thickness and bottom temperature, as summarized in Table 1). The range of variation of each parameter has been carefully defined with the aim to represent the future uncertainties, keeping in mind that the lake water budget will be eventually resolved in our modeling.

With this work, we have shown the relevance of the SA methods to better understand model functioning and parameter influences. The added information brought by the analysis of SI time series has been also demonstrated here. The time varying sensitivities of the input parameters was clearly demonstrated, from the dominance of the snow/ice ones in frozen conditions, to the relative contribution of the other ones for free water according to lake depth, that is, larger sensitivity of lake depth and light extinction coefficient for shallow lakes whereas water albedo and thermocline relaxation constant are more influential for deeper lakes.

The analysis was limited to LWST and surface turbulent fluxes because these are the main variables seen by the atmosphere in a climate model. It has been fully informative to better understand the functioning of FLake model, its parameters and interactions. It allows us to start its implementation in ORCHIDEE with confidence and to identify the databases that will be used for its calibration. The altimetry and most recent multispectral space missions will be given first priority for that purpose.

## Data Availability Statement

The Flake model used in this study was downloaded from <http://www.flake.igb-berlin.de/site/download>. The codes used to generate the results of the sensitivity analysis can be downloaded from <https://doi.org/10.5281/zenodo.3700135>

## References

- Balsamo, G. (2010). Deriving an effective lake depth from satellite lake surface temperature data: A feasibility study with MODIS data. *Boreal Environment Research*, 15(2), 178–190.
- Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T., & Potes, M. (2012). On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus A: Dynamic Meteorology and Oceanography*, 64(1), 15829. <https://doi.org/10.3402/tellusa.v64i0.15829>
- Bastviken, D., Cole, J., Pace, M., & Tranvik, L. (2004). Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate: Lake Methane Emissions. *Global Biogeochemical Cycles*, 18(4). <https://doi.org/10.1029/2004gb002238>
- Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. *Surveys in Geophysics*, 37(2), 307–337. <https://doi.org/10.1007/s10712-015-9346-y>
- Bonan, G. B. (1995). Sensitivity of a GCM simulation to inclusion of inland water surfaces. *Journal of Climate*, 8(11), 2691–2704. [https://doi.org/10.1175/1520-0442\(1995\)008<2691:SOAGST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<2691:SOAGST>2.0.CO;2)
- Chuang, H.-Y., & Sousounis, P. J. (2003). The impact of the prevailing synoptic situation on the lake-aggregate effect. *Monthly Weather Review*, 131(5), 990–1010. [https://doi.org/10.1175/1520-0493\(2003\)131<0990:TIOPTS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0990:TIOPTS>2.0.CO;2)
- Créteaux, J.-F., Arsen, A., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., et al. (2011). SOLS: A lake database to monitor in the near real time water level and storage variations from remote sensing data. *Advances in Space Research*, 47(9), 1497–1507. <https://doi.org/10.1016/j.asr.2011.01.004>
- d'Orgeval, T., Polcher, J., & de Rosnay, P. (2008). Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes. *Hydrology and Earth System Sciences*, 12(6), 1387–1401. <https://doi.org/10.5194/hess-12-1387-2008>
- Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P., Mironov, D., & Schar, C. (2010). An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal Environment Research*, 15(2), 100–112.
- Ebert, E. E., & Curry, J. A. (1993). An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions. *Journal of Geophysical Research*, 98(C6), 10085. <https://doi.org/10.1029/93jc00656>
- Saltelli, A. (Ed.). (2008). *Global sensitivity analysis: The primer* (OCLC: ocn180852094). John Wiley.
- Goloso, S., & Kirillin, G. (2010). A parameterized model of heat storage by lake sediments. *Environmental Modelling & Software*, 25(6), 793–801. <https://doi.org/10.1016/j.envsoft.2010.01.002>
- Goloso, S., Zverev, I., Shipunova, E., & Terzhevik, A. (2018). Modified parameterization of the vertical water temperature profile in the FLake model. *Tellus A: Dynamic Meteorology and Oceanography*, 70(1), 1–7. <https://doi.org/10.1080/16000870.2018.1441247>
- Guerrero, J.-L., Pernica, P., Wheeler, H., Mackay, M., & Spence, C. (2017). Parameter sensitivity analysis of a 1-D cold region lake model for land-surface schemes. *Hydrology and Earth System Sciences*, 21(12), 6345–6362. <https://doi.org/10.5194/hess-21-6345-2017>
- Heiskanen, J. J., Mammarella, I., Ojala, A., Stepanenko, V., Erkkilä, K. M., Miettinen, H., et al. (2015). Effects of water clarity on lake stratification and lake-atmosphere heat exchange. *Journal of Geophysical Research - D: Atmospheres*, 120(15), 7412–7428. <https://doi.org/10.1002/2014JD022938>

## Acknowledgments

This work was essentially 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'Énergie Atomique et aux énergies alternatives) is also acknowledged for funding part of A. Bernus PhD grant as well as the H2020 MULTIPLY project (Grant No. 687320) funding N. Raoult. 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. The authors are also grateful to Dr. D. Mironov and G. Kirillin for valuable comments and helpful discussions. The reviewers are also acknowledged for their constructive comments.

- ILEC. (1999). *World Lake database*. International Lake Environment Committee Foundation. Retrieved from <http://wldb.ilec.or.jp>
- Laird, N., Sobash, R., & Hodas, N. (2009). The frequency and characteristics of lake-effect precipitation events associated with the New York State Finger Lakes. *Journal of Applied Meteorology and Climatology*, 48(4), 873–886. <https://doi.org/10.1175/2008JAMC2054.1>
- Lamboni, M., Monod, H., & Makowski, D. (2011). Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering & System Safety*, 96(4), 450–459. <https://doi.org/10.1016/j.ress.2010.12.002>
- Launiainen, J., & Cheng, B. (1998). Modelling of ice thermodynamics in natural water bodies. *Cold Regions Science and Technology*, 27(3), 153–178. [https://doi.org/10.1016/S0165-232X\(98\)00009-3](https://doi.org/10.1016/S0165-232X(98)00009-3)
- Layden, A., MacCallum, S. N., & Merchant, C. J. (2016). Determining lake surface water temperatures worldwide using a tuned one-dimensional lake model (FLake, v1). *Geoscientific Model Development*, 9(6), 2167–2189. <https://doi.org/10.5194/gmd-9-2167-2016>
- Le Moigne, P., Colin, J., & Decharme, B. (2016). Impact of lake surface temperatures simulated by the FLake scheme in the CNRM-CM5 climate model. *Tellus A: Dynamic Meteorology and Oceanography*, 68(1), 31274. <https://doi.org/10.3402/tellusa.v68.31274>
- Li, J., Duan, Q. Y., Gong, W., Ye, A., Dai, Y., Miao, C., et al. (2013). Assessing parameter importance of the Common Land Model based on qualitative and quantitative sensitivity analysis. *Hydrology and Earth System Sciences*, 17(8), 3279–3293. <https://doi.org/10.5194/hess-17-3279-2013>
- Ljungemyr, P., Gustafsson, N., & Omstedt, A. (1996). Parameterization of lake thermodynamics in a high-resolution weather forecasting model. *Tellus*, 48(5), 608–621. <https://doi.org/10.1034/j.1600-0870.1996.t01-4-00002.x>
- Martynov, A., Sushama, L., & Laprise, R. (2010). Simulation of temperate freezing lakes by one-dimensional lake models: Performance assessment for interactive coupling with regional climate models. *Boreal Environment Research*, 15(1), 143–164.
- McMahon, A., & Moore, R. D. (2017). Influence of turbidity and aeration on the albedo of mountain streams. *Hydrological Processes*, 31(25), 4477–4491. <https://doi.org/10.1002/hyp.11370>
- Messenger, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7(1). <https://doi.org/10.1038/ncomms13603>
- Mironov, D. (2008). Parameterization of lakes in numerical weather prediction. In Description of a lake model (Technical Report No. 11). Deutscher Wetterdienst.
- Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., & Terzhevik, A. (2010). Implementation of the lake parameterization scheme FLake into the numerical weather prediction model COSMOS. *Boreal Environment Research Publishing Board*, 15, 218–230.
- Mironov, D., Ritter, B., Schulz, J.-P., Buchhold, M., Lange, M., & MacHulskaya, E. (2012). Parameterization of sea and lake ice in numerical weather prediction models of the German weather service. *Tellus A: Dynamic Meteorology and Oceanography*, 64(1), 17330. <https://doi.org/10.3402/tellusa.v64i0.17330>
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Poole, H. H., & Atkins, W. R. G. (1929). Photo-electric measurements of submarine illumination throughout the year. *Journal of the Marine Biological Association of the United Kingdom*, 16(1), 297–324. <https://doi.org/10.1017/S0025315400029829>
- Rooney, G. G., & Bornemann, F. J. (2013). The performance of FLake in the met office unified model. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1), 21363. <https://doi.org/10.3402/tellusa.v65i0.21363>
- Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X., & de Gonçalves, L. G. G. (2012). A fully multiple-criteria implementation of the Sobol' method for parameter sensitivity analysis. *Journal of Geophysical Research*, 117(D7). <https://doi.org/10.1029/2011JD016355>
- Salgado, R., & Le Moigne, P. (2010). Coupling of the FLake model to the surfex externalized surface model. *Boreal Environ Research Publishing Board*, 15, 231–244.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2), 259–270. <https://doi.org/10.1016/j.cpc.2009.09.018>
- Sandström, G. (2003). Impact of lakes and wetlands on boreal climate. *Journal of Geophysical Research*, 108(D16). <https://doi.org/10.1029/2002JD002597>
- Sobol, I. (1990). Quasi-Monte Carlo methods. *Progress in Nuclear Energy*, 24(1–3), 55–61. [https://doi.org/10.1016/0149-1970\(90\)90022-W](https://doi.org/10.1016/0149-1970(90)90022-W)
- Song, Y., Semazzi, F. H. M., Xie, L., & Ogallo, L. J. (2004). A coupled regional climate model for the Lake Victoria basin of East Africa. *International Journal of Climatology*, 24(1), 57–75. <https://doi.org/10.1002/joc.983>
- Stepanenko, V. (2005). Numerical simulation of the thermal regime of shallow water bodies. *Computational Technologies*, 10(1), 100–106.
- Tang, Y., Reed, P., Wagener, T., & van Werkhoven, K. (2007). Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences*, 11(2), 793–817. <https://doi.org/10.5194/hess-11-793-2007>
- Telmer, K. H., & Costa, M. P. F. (2007). SAR-based estimates of the size distribution of lakes in Brazil and Canada: A tool for investigating carbon in lakes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 17(3), 289–304. <https://doi.org/10.1002/aqc.854>
- Thiery, W., Davin, E. L., Panitz, H.-J., Demuzere, M., Lhermitte, S., & van Lipzig, N. (2015). The impact of the African great lakes on the regional climate. *Journal of Climate*, 28(10), 4061–4085. <https://doi.org/10.1175/JCLI-D-14-00565.1>
- Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18), 6396–6402. <https://doi.org/10.1002/2014GL060641>
- Voros, M., Istvnovics, V., & Weidinger, T. (2010). Applicability of the FLake model to lake Balaton. *Boreal Environment Research*, 15(2), 245–254.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50(9), 7505–7514. <https://doi.org/10.1002/2014WR015638>
- West, W. E., Creamer, K. P., & Jones, S. E. (2016). Productivity and depth regulate lake contributions to atmospheric methane. *Limnology & Oceanography*, 61(S1), S51–S61. <https://doi.org/10.1002/lno.10247>
- Woolway, R. I., & Merchant, C. J. (2019). Worldwide alteration of lake mixing regimes in response to climate change. *Nature Geoscience*, 12(4), 271–276. <https://doi.org/10.1038/s41561-019-0322-x>
- Xu, J., Shu, H., Jiang, H., & Dong, L. (2012). Sobol sensitivity analysis of parameters in the common land model for simulation of water and energy fluxes. *Earth Science Informatics*, 5(3–4), 167–179. <https://doi.org/10.1007/s12145-012-0105-z>
- Zhang, C., Chu, J., & Fu, G. (2013). Sobol's sensitivity analysis for a distributed hydrological model of Yichun River basin, China. *Journal of Hydrology*, 480, 58–68. <https://doi.org/10.1016/j.jhydrol.2012.12.005>
- Zheng, H., Yang, Z. L., Lin, P., Wei, J., Wu, W. Y., Li, L., et al. (2019). On the sensitivity of the precipitation partitioning into evapotranspiration and runoff in land surface parameterizations. *Water Resources Research*, 55(1), 95–111. <https://doi.org/10.1029/2017wr022236>



## 4 - Paramétrisation des surfaces et des profondeurs des lacs

### Sommaire

---

<b>4.1 Répartition des lacs à l'échelle globale . . . .</b>	<b>85</b>
4.1.1 Occupation des sols dans ORCHIDEE . . . .	85
4.1.2 Estimation des fractions de lacs . . . . .	86
<b>4.2 Spatialisation des profondeurs . . . . .</b>	<b>88</b>
4.2.1 HydroLAKES . . . . .	88
4.2.2 Suivi des séries temporelles des profondeurs .	89
4.2.3 Agrégation à la résolution des cartes de PFT	91

---

### 4.1 . Répartition des lacs à l'échelle globale

#### 4.1.1 . Occupation des sols dans ORCHIDEE

Pour simuler le bilan d'énergie des surfaces continentales, la première étape consiste à rendre compte de l'hétérogénéité des surfaces en les différenciant et les classant. On utilise pour cela des cartes d'occupation des sols disponibles dans la communauté dont la résolution varie selon les produits entre 10 m et quelques centaines de kilomètres. Par exemple, au niveau européen, nous pouvons citer le produit CORINE à une résolution de 30 m construit à partir des images satellitaires Landsat 7, SPOT-4 et IRS LISS.

Dans notre cas, nous avons besoin de cartes à l'échelle globale. Dans les modèles de surfaces continentales COSMO et SURFEX, le produit ECOCLIMAP-II (Faroux et al., 2013) est utilisé. Il différencie 270 classes représentées avec certains paramètres physiques associés comme l'albédo ou le LAI à une résolution kilométrique. Dans le modèle ORCHIDEE, le produit ESA-CCI (ESA, 2017) est utilisé pour créer les cartes de PFT.

Ce produit d'occupation des sols a une résolution temporelle annuelle (entre 1992-2015) et spatiale de 300 m et contient 22 classes. Les instruments satellitaires utilisés sont différents selon les années traitées. Les observations d'AVHRR sont utilisées pour la période 1992-1999, de SPOT-VGT pour 1999-2013, de PROBA-V pour 2013-2015 et de MERIS pour 2003 et 2012. Certaines incertitudes sont liées à des fréquences de passage plus faibles sur certaines régions ou encore à la difficulté de séparer correctement certaines classes.

La figure 4.1 montre la carte d'occupation des sols du produit ESA-CCI pour l'année 2015. Le code couleur utilisé pour faire la figure est décrit dans la Table 4.1.1. Grâce à ces cartes, on peut par exemple voir l'extension des surfaces des

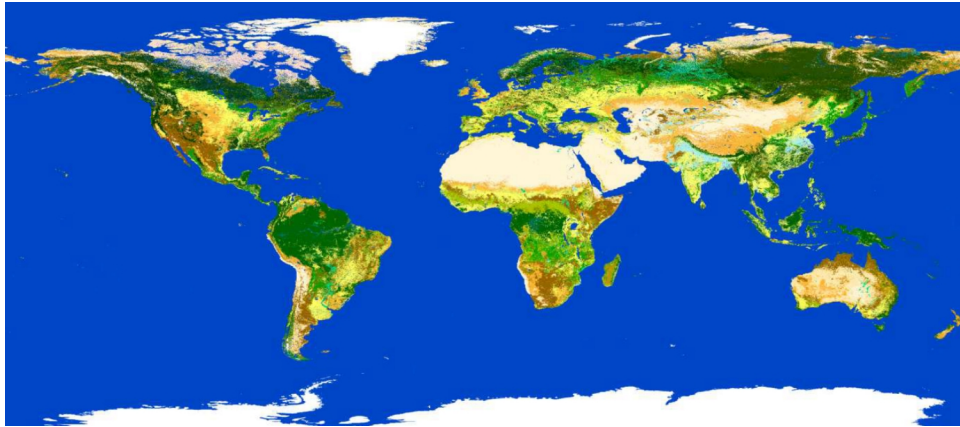


FIGURE 4.1 – Carte d’occupation des sols de l’année 2015 du produit ESA-CCI à la résolution de 300 m. Les couleurs présentes sont associées à un type de surface et sont décrites par la table 4.1.1

viles qui représente 0,057% de la surface de la Terre en 1992 et 0,12% en 2015. Cette carte distingue depuis récemment les eaux continentales et océaniques et les surfaces d’eau temporaires et pérennes (Lamarche et al., 2017). Cependant, les lacs ne sont pas distincts des rivières par exemple.

Pour estimer les PFT d’ORCHIDEE, une table de conversion permet de convertir les classes ESA-CCI en PFT, elle est accessible sur le site <https://orchid.as.lscce.ipsl.fr/dev/lccci/tools.php>. La carte de Koppen-Geiger permet de séparer cinq climats différents et de raffiner le découpage des classes CCI en fonction du climat du pixel. De même pour les prairies et les cultures, la base de données développée par Still et al. (2003) permet de différencier les deux types de photosynthèse C3 et C4. Poulter et al. (2011) détaille ce processus depuis la carte CCI-ESA jusqu’aux PFT finaux. La dernière étape est une agrégation des PFT à une certaine résolution (ici  $0.5^\circ$ ) pour obtenir leur fraction dans chaque maille de la grille.

#### 4.1.2 . Estimation des fractions de lacs

La carte ESA-CCI ne distingue pas les lacs des autres surfaces d’eau libre comme les rivières. Il faut donc une autre base de données globale pour les lacs. Il en existe plusieurs. Lehner and Döll (2004) ont développé une première base de données GLWD conséquente avec plus de 250000 lacs de plus de 10 ha. Celle-ci est construite sur sept autres bases de données antérieures (globales ou régionales) et permet de différencier les différentes surfaces d’eau (lacs, rivières, ...). Verpoorter et al. (2014) ont utilisé des images du satellite Landsat 7 pour créer la base de données GLOWABO. La base de données dénombre 117 millions de lacs d’une taille supérieure à  $0.002 \text{ km}^2$ . GLOWABO dénombre moins de petits lacs que des interpolations statistiques de la forme  $c \cdot A^{-b}$  (avec c et b des constantes et A la
























Identifiant	Classe de la surface	Couleur	Identifiant	Classe de la surface	Couleur
0	Pas de données		120	Arbustes	
10	Culture pluviale		130	Prairie	
20	Culture irriguée		140	Mousses et Lichens	
30	Pixel mixte culture >50%, végétation naturelle <50%)		150	Végétation clairsemée (<15%)	
40	Pixel mixte culture <50%, végétation naturelle >50%)		160	Forêt inondée (eau douce)	
50	Forêt de Feuillus à feuillage persistant (>15%)		170	Forêt inondée (eau saline)	
60	Forêt de Feuillus mixte à feuillage temporaire (>15%)		180	Herbacées/Arbuste inondés	
70	Forêt de conifères à aiguilles persistantes (>15%)		190	Zone urbaine	
80	Forêt de conifères à aiguilles temporaires (>15%)		200	Sol nu	
90	Forêt mixte (feuillus et conifères)		210	Eau libre	
100	Pixel mixte (Arbres et arbustes >50%, prairies <50%)		220	Neige et glace permanentes	
110	Pixel mixte (Arbres et arbustes <50%, prairies >50%)				

TABLE 4.1 – Classification du produit ESA-CCI landcover



surface du lac) comme le suggérait l'étude de [Seekell et al. \(2013\)](#). Par la suite, la base de données de [Pekel et al. \(2016\)](#) a permis d'avoir des séries temporelles des surfaces de lacs au cours de 32 années avec un intervalle de 5 ans et des images Landsat d'une résolution de 30 m. Ils ont notamment mis en évidence que les surfaces d'eau permanentes sont plus importantes au cours du temps. Cela est principalement lié au remplissage des réservoirs et dans une moindre mesure au réchauffement climatique. [Sheng et al. \(2016\)](#) crée un nouvel algorithme qu'ils appliquent sur les images Landsat 8 pour créer une nouvelle base de données. Cet algorithme est notamment repris dans le cadre de la mission SWOT pour détecter les lacs (voir description dans la prochaine section). HydroLAKES est une autre base de données qui recense 1.43 millions de lacs de plus de 0.1 km<sup>2</sup> qui se base principalement sur GLWD, mais aussi sur d'autres bases de données globales ou régionales. L'originalité de cette étude est l'estimation des profondeurs de chaque lac qui sera expliqué dans le prochain paragraphe. La base de données de [Kourzeneva \(2010\)](#) est quant à elle construite à partir de base de données préexistantes, mais de manière à être en cohérence avec ECOCLIMAP. Plus récemment cette base de donnée a été amélioré par [Choulga et al. \(2019\)](#).

Dans notre cas, nous avons choisi la base de données HydroLAKES, car celle-ci est cohérente avec HydroSHEDS, une base de données du réseau de drainage et des rivières, qui est utilisée dans le système de routage de l'eau d'ORCHIDEE. De plus, seule cette base de données avec celle de [Kourzeneva \(2010\)](#) fournit la profondeur moyenne des lacs qui est un paramètre primordial pour FLake.

## 4.2 . Spatialisation des profondeurs

### 4.2.1 . HydroLAKES

HydroLAKES fournit une estimation de la profondeur de plus d'un million de lacs avec des modèles statistiques. Une première étape est de générer les surfaces terrestres à l'aide du logiciel EarthEnv-DEM90 ([Robinson et al., 2014](#)) à une résolution d'environ 90 m pour ensuite obtenir les pentes en chaque point. Plusieurs modèles statistiques ont été testés avec des variables explicatives différentes. Le modèle final (voir Table 4.2) sépare quatre catégories de lacs et utilise la surface du lac notée  $A$  et la pente moyenne  $S_{100}$  autour du contour du lac dans un rayon de 100 m. Le modèle a été entraîné sur 7049 lacs et validé sur 5101 lacs pour obtenir les coefficients des équations pour chaque type de lac avec un écart-type résiduel du modèle entraîné donné dans la Table 4.2.

Le modèle prédit le logarithme de la profondeur pour éviter une certaine hétéroscédasticité de la prédiction. Pour mémoire, l'hétéroscédasticité mesure la variabilité de la variance des erreurs selon les différentes variables stochastiques indépendantes. Dans le modèle linéaire table 4.2, minimiser l'hétéroscédasticité signifie que la variance de l'erreur doit être constante quelles que soient les valeurs de la surface  $A$  et de la pente moyenne  $S_{100}$ . Le passage du logarithme à la profondeur

Catégorie de lac (surface en $km^2$ )	Équation du modèle	Ecart type
0.1-1	$\log_{10}(D) = 0.3826 + 0.1512\log_{10}(A)$ $+ 0.4820 \cdot \log_{10}(S_{100})$	0.0678
1-10	$\log_{10}(D) = 0.1801 + 0.2985 \cdot \log_{10}(A)$ $+ 0.8473 \cdot \log_{10}(S_{100})$	0.0689
10-100	$\log_{10}(D) = 0.0379 + 0.2445 \cdot \log_{10}(A)$ $+ 1.1517 \cdot \log_{10}(S_{100})$	0.0692
100-500	$\log_{10}(D) = 0.0123 + 0.2664 \cdot \log_{10}(A)$ $+ 1.1474 \cdot \log_{10}(S_{100})$	0.1094

TABLE 4.2 – Équations du modèle statistique utilisé dans la base de données HydroLAKES

donne un biais inhérent. Il faut donc ensuite débiaiser la profondeur comme les travaux de [Heathcote et al. \(2015\)](#) le suggère avec l'équation ???. Dans la formule ??,  $s$  représente l'écart-type qui provient du modèle entraîné (voir table 4.2)

$$D_{biascorrected} = 10^{\log_{10}(D_{model})} e^{-2.65s^2} \quad (4.1)$$

#### 4.2.2 . Suivi des séries temporelles des profondeurs

La profondeur est un paramètre qui peut varier très fortement temporellement. Dans le chapitre précédent, cette variabilité de la variation saisonnière a été estimée pour calculer l'intervalle de la profondeur pour l'analyse de sensibilité en utilisant la base de données Hydroweb ([Créteaux et al., 2011](#)). Cette base de données permet d'obtenir la série temporelle des hauteurs d'eau de 300 lacs dans le monde. Différents satellites sont utilisés selon la période envisagée en utilisant le concept de l'altimétrie spatiale. Par exemple, on peut voir dans la base Hydroweb que sur la période 1992 à 2021, la variation du lac Victoria est de presque  $3m$  comme pour le lac Malawi entre le niveau d'eau le plus bas et le plus haut. Pour les lacs de faible profondeur, une telle variation peut changer de façon importante le bilan d'énergie. Prendre en compte cette évolution n'est pas négligeable dans la paramétrisation. Un suivi satellitaire des hauteurs d'eau, comme le permet l'altimétrie spatiale, peut pallier à ce problème.

##### 4.2.2.1 Principe de l'altimétrie nadir

Le principe repose sur la mesure du temps entre l'émission d'une onde radar et la réception de celle-ci par la surface considérée. Le point qui réfléchit l'onde est au nadir du satellite, c'est-à-dire le point de la surface à la verticale du satellite.

De plus, la connaissance de l'orbite du satellite permet de connaître la distance entre le satellite et la surface de référence terrestre que l'on nomme géodésique de référence. La différence entre ces deux distances permet de calculer l'élévation de la surface d'eau après avoir pris en compte un certain nombre de corrections instrumentales liées à la propagation des ondes électromagnétiques dans l'atmosphère et liées à la topographie environnante (Cretaux et al., 2017). Par exemple, les données d'Hydroweb proviennent notamment des satellites Topex/Poseidon (1992-2006), Jason-1 (2001-2013) et 2 (2008-2019), ENVISAT (2002-2012) et GFO qui utilisent l'altimétrie nadir. D'autres missions satellitaires nous donnent des informations sur l'élévation des surfaces d'eau, on peut distinguer deux modes d'acquisition. Les altimètres nadir sur les satellites comme ERS-1 et 2 ou SARAL fonctionnent seulement avec le mode LRM (Low Resolution Mode) qui traite l'onde reçue sans tenir compte de l'effet doppler. Plus récemment, les altimètres nadir des satellites comme Cryosat-2 ou Sentinel-3 fonctionnent avec le mode SAR qui permet d'émettre plus d'impulsions d'onde pour utiliser l'effet doppler. Ainsi, comme la vitesse relative de l'onde par rapport au satellite change selon l'angle d'émission, la fréquence de l'onde change aussi, ce qui permet d'augmenter la résolution dans la trajectoire du satellite. Les fréquences radar utilisées par les altimètres sont présentes dans les bandes suivantes : Ku (12 à 18 GHz) et plus récemment Ka (26.5 à 40 GHz). L'une des limitations de ces technologies est la localisation de la mesure sur un point donné.

#### 4.2.2.2 La mission SWOT

Dans l'optique de suivre l'évolution des surfaces d'eau continentale au niveau global, la mission WATER fut proposée par Alsdorf et al. (2007) pour utiliser les instruments des missions satellitaires océanographiques. La mission a été pensée de manière conjointe avec la communauté océanographique, le projet final a été renommé SWOT et conduit par la NASA et le CNES. L'Agence Spatiale Canadienne (ASC) et l'Agence Spatiale du Royaume-Uni (UKSA) sont aussi impliquées dans le projet. La mission SWOT a pour objectif de faire des mesures à fine échelle de la topographie des océans pour mieux comprendre des processus méso et submésos-échelles et de faire des mesures des hauteurs d'eau des surfaces d'eau terrestres pour mieux comprendre leur distribution spatiale et temporelle et aussi les débits entrants et sortants des lacs et des rivières.

Pour suivre l'évolution des hauteurs d'eau au niveau global, la mission SWOT utilise une autre technologie que l'altimétrie nadir pour avoir des mesures spatialement plus étendues des surface d'eau continentale (Biancamaria et al., 2016). La surface visualisée par les instruments se nomme la fauchée. L'instrument utilisé KaRin est un interféromètre bi-statique à faible incidence fonctionnant en bande Ka qui permettra de cartographier les hauteurs d'eau sur les océans et les continents (Fjortoft et al., 2014). La dénomination bistatique fait référence à la réception des ondes qui se fait sur une antenne différente que l'émission comme on le voit sur

la Figure 4.2. L'interféromètre se fonde sur le même principe qu'un altimètre classique. Cependant, l'onde reçue concerne une zone étendue et la mesure représente donc une intégration. Le post-traitement est donc plus compliqué et des techniques comme la transformée de Fourier permettent de remonter à l'information voulue. Le cycle orbital de SWOT sera de 21 jours. La résolution horizontale sera de quelques mètres pour les étendues d'eau plus grande de  $0.06 \text{ km}^2$  (Biancamaria et al., 2016). Après agrégation des mesures sur des surfaces d'eau libre supérieures à  $1 \text{ km}^2$ , l'instrument pourra fournir le niveau de l'eau avec une précision de l'ordre de 10 cm (Sachse et al., 2014).

Dans notre cas, les données récoltées par le satellite nous permettront de suivre l'évolution des hauteurs d'eau des lacs. Ces données pourront améliorer la paramétrisation des modèles comme FLake comme nous l'avons montré dans l'article Ottlé et al. (2020) qui figure en Annexe 5.4. Dans cette étude, nous avons étudié quatre lacs réservoirs de la Seine. Nous avons comparé une simulation de référence sur la période 2016-2017 où la profondeur est actualisée chaque jour, une simulation avec une profondeur constante égale à la hauteur maximale du réservoir et une simulation avec les hauteurs actualisées avec les jours de passage de SWOT et bruitées avec un simulateur d'erreur du satellite. Ainsi, la racine de l'écart quadratique moyen de la température de surface varie entre  $1.2 \text{ }^\circ\text{C}$  et  $0.9 \text{ }^\circ\text{C}$  dans la simulation sans les observations SWOT et avec les observations SWOT, on estime que cette erreur varie entre  $0.2 \text{ }^\circ\text{C}$  et  $0.5 \text{ }^\circ\text{C}$  selon les lacs.

#### 4.2.3 . Agrégation à la résolution des cartes de PFT

À cette étape, nous avons seulement une base de données des surfaces et des profondeurs non dynamiques, il faut néanmoins l'agréger à la résolution des cartes de PFT.

##### 4.2.3.1 Étude des distributions des profondeurs à $0.5^\circ$

Les forçages utilisés ont une résolution entre  $0.25^\circ$  et  $0.5^\circ$ . Une première étape est d'étudier la distribution des profondeurs à cette résolution dans chaque pixel. On agrège les profondeurs à ces deux résolutions en prenant la moyenne pour chaque tranche de profondeur. Comme le nombre de lacs profonds est moins élevé, le pas de discrétisation augmente avec la profondeur comme indiqué sur la Table 4.3.

La carte dénombrant les lacs traités dans chaque point de grille est tracée sur la figure 4.3 pour la résolution de  $0.5^\circ$ .

Le nombre maximal de lacs dans un point de grille atteint environ 800 et se situe au Canada. Dans les zones où la fraction de lacs est élevée comme au Canada ou dans les pays scandinaves, de nombreux pixels ORCHIDEE peuvent avoir à traiter plus de 100 lacs. L'agrégation peut donc avoir un réel impact sur la paramétrisation de la profondeur.

La figure 4.4 représente la carte des fractions de lacs agrégées à  $0.5$  degré. Les

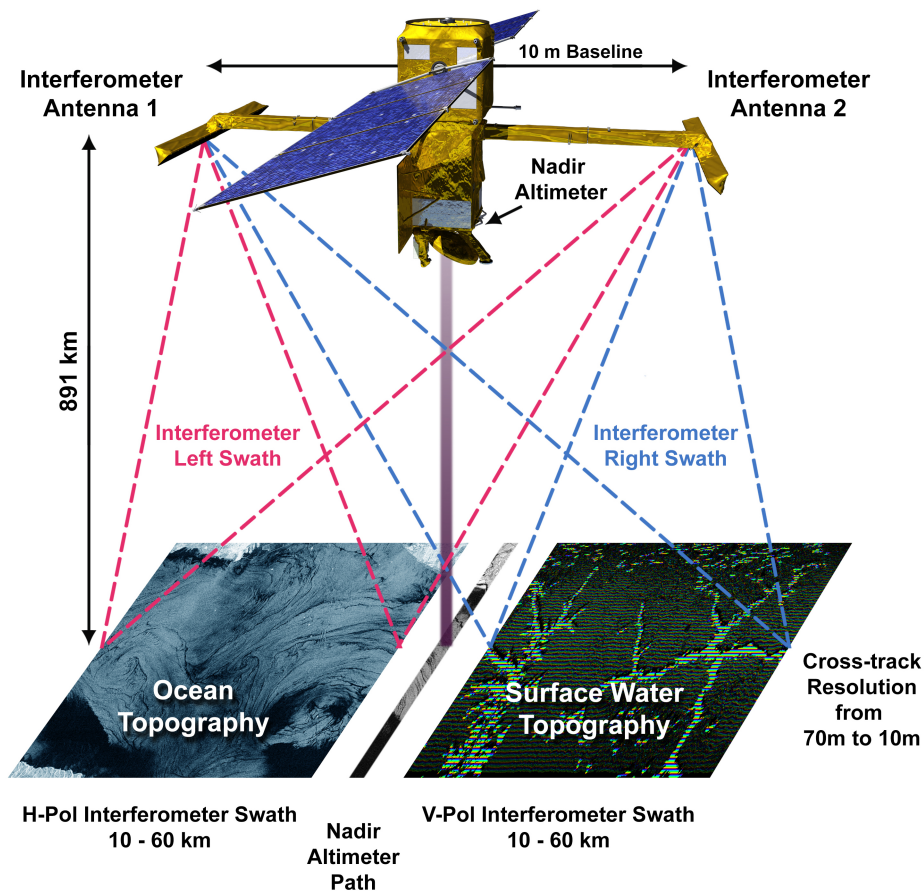


FIGURE 4.2 – Schéma représentant le satellite SWOT en orbite (source [Fu et al. \(2012\)](#))

distributions des fractions et du nombre de lacs sont tracées aussi sur une région particulière (au-dessus des grands lacs américains) représentée par un rectangle rouge sur la carte globale. Cette zone a été choisie pour sa fraction élevée de lacs. En tout, 48 pixels ont été tracés pour mieux se rendre compte de la variabilité spatiale de la distribution des surfaces et des profondeurs. Il y a un très grand nombre de lacs de faible profondeur avec un mode des distributions des profondeurs qui se situe autour de 3 et 4 mètres. Cependant, ils représentent une fraction du même ordre de grandeur qu'un lac plus profond au-delà de 10 m qui est représenté par un pic important sur la distribution des fractions et qui n'apparaît pas sur la distribution du nombre de lacs. Ces distributions de fraction sont plutôt bi-modales avec un mode étendu pour les profondeurs en dessous de 5 m et un mode étroit pour les profondeurs de plus de 10 m.

Si nous choisissons un seul type de lac dans la maille, le choix des indicateurs

Intervalle de profondeur (m)	Discrétisation (m)
0-50 m	0.1
50-100 m	0.2
100-200 m	1
200-500 m	10

TABLE 4.3 – Discrétisation de la carte des profondeurs

statistiques pour agréger la profondeur peut avoir un réel impact sur la profondeur agrégée comme on le montrera dans le chapitre suivant. Nous avons notamment comparé la moyenne arithmétique, harmonique, quadratique et la médiane des profondeurs. Ces façons d'agréger les profondeurs peuvent donner lieu à plusieurs mètres d'écart sur la valeur finale retenue.

#### 4.2.3.2 Choix d'agrégation

Dans un même pixel, des lacs avec des profondeurs très différentes peuvent être présents. Nous avons montré dans l'analyse de sensibilité que des lacs très et peu profonds ont des comportements distincts. Il semble donc légitime de les séparer en plusieurs catégories. L'analyse de sensibilité montre aussi que la température est malgré tout sensible à l'incertitude de la profondeur à l'intérieur d'une même classe et d'autant plus que le lac est peu profond.

Ainsi les différents indicateurs statistiques pour agréger les lacs pourraient avoir un réel impact sur les températures. Les indicateurs statistiques utilisés sont pondérés par les surfaces, car ce sont elles qui jouent un rôle important dans le bilan d'énergie et non le nombre de lacs. Ainsi, nous avons considéré la moyenne arithmétique, car c'est l'indicateur le plus intuitif et il ne privilégie pas un type de lac en particulier. Nous avons aussi pris la médiane qui permet d'atténuer l'effet des lacs marginaux. Nous avons considéré la moyenne harmonique qui privilégie les petites profondeurs si nous considérons seulement un type de lac dans la maille. En effet, le modèle FLake est moins sensible aux grandes profondeurs et donc un lac de profondeur moyenne (par exemple 25 m) va avoir un comportement pour ce qui concerne la surface très similaire à un lac de 50 m. Ainsi, pour avoir un comportement moyen quand on agrège les lacs de faibles et grandes profondeurs,

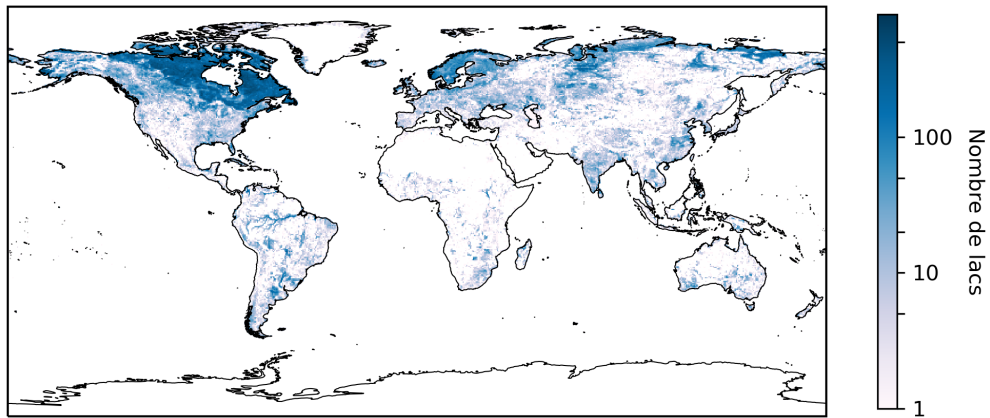


FIGURE 4.3 – Carte représentant le nombre de lacs traités par point de grille à une résolution de  $0.5^\circ$ .

on peut donner un poids plus important pour les lacs peu profonds. Nous avons aussi essayé la moyenne quadratique qui a un effet contraire en favorisant les lacs les plus profonds. Nous avons aussi considéré trois types de lac en se référant au nombre de catégories choisies dans l'analyse de sensibilité. Nous avons aussi préparé une autre stratégie pour agréger les profondeurs des lacs. Nous avons lissé les distributions de profondeur et sélectionné les pics les plus importants. Cependant, nous ne l'avons pas testé dans la suite du travail.

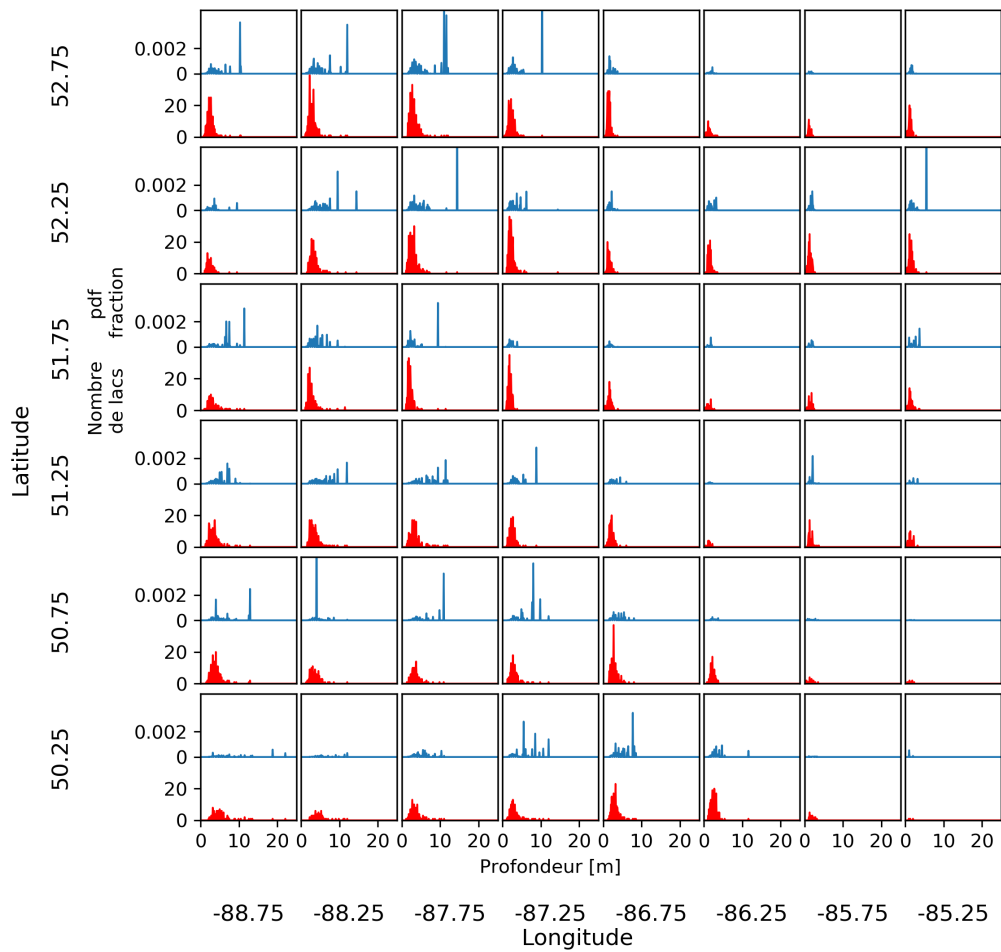
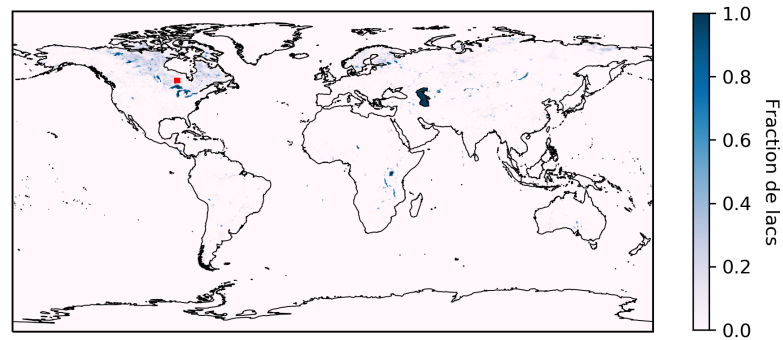


FIGURE 4.4 – Représentation des fractions des lacs à une résolution de  $0.5^\circ$  en haut avec un zoom représenté par un rectangle rouge sur la carte globale. Le tableau représente les distributions des profondeurs en bleu et du nombre de lacs en rouge pour chaque pixel de la région considérée. Chaque ligne représente une longitude et chaque colonne une latitude.





## 5 - Evaluation du couplage ORCHIDEE FLake

### Sommaire

---

5.1	Introduction . . . . .	97
5.2	Méthode . . . . .	97
5.3	Synthèse des résultats . . . . .	98
5.4	Article 2 . . . . .	98

---

#### 5.1 . Introduction

Dans ce chapitre, nous présentons la validation du couplage entre les modèles ORCHIDEE et FLake. Nous avons confronté les simulations avec des observations satellitaires de la température de surface et les avons évaluées avec la Racine de l'erreur quadratique moyenne (Root mean square error) (RMSE). Avec les résultats de l'analyse de sensibilité qui confirme l'influence de la profondeur, nous nous sommes focalisés sur la stratégie de ce paramètre. Nous avons notamment évalué la nouvelle stratégie de paramétrisation qui consiste à séparer la fraction de lac en trois classes en fonction de leur profondeur (inférieure à 5 m, entre 5 et 25 m, supérieure à 25 m).

#### 5.2 . Méthode

Le modèle ORCHIDEE est couplé avec FLake avec les améliorations présentées dans la section 2.4.6, notamment sur la représentation de la couche de neige qui suit les recommandations de [Semmler et al. \(2012\)](#) et sur l'évolution sur le coefficient de forme de la thermocline proposé par [Salgado and Le Moigne \(2010\)](#).

La préparation du modèle pour réaliser les simulations consiste à agréger les différentes cartes telles que la carte d'occupation des sols ESA-CCI et la carte des profondeurs des lacs d'HydroLAKES à la résolution des forçages utilisés (0.25 et 0.5 degrés).

Plusieurs cartes des profondeurs ont été générées pour évaluer plusieurs stratégies (5 au total : 4 cartes avec un type de lac par maille et une carte avec 3 types). Nous avons testé plusieurs indicateurs statistiques pour agréger les lacs : la moyenne, la médiane, les moyennes harmonique et quadratique. La médiane était le meilleur indicateur sur les lacs observés, nous l'avons donc sélectionnée pour la carte avec 3 types de lacs.

Nous avons aussi utilisé 5 forçages (WFDEI, CRUJRA, CRUNCEP, E2OFD et ERA5) pour étudier la sensibilité du modèle au forçage. ERA5 et E2OFD sont des forçages à 0.25 degrés alors que les autres sont à 0.5 degrés. La Table 1 de l'article

synthétise les caractéristiques de chaque simulation.

Les simulations ont été réalisées avec 25 années de spin-up et sont exploitables entre 2000 et 2016. La base de données GloboLakes est utilisée pour valider les températures de surface et la base de données Global Lake and River Ice Phenology Database pour la durée de gel simulée. L'ensemble de ces résultats ont fait l'objet d'un article en cours de révision pour la revue Geoscientific Model Development (GMD) et qui figure dans le chapitre suivant.

### 5.3 . Synthèse des résultats

L'évaluation des résultats est sensible au forçage météorologique. En comparant le RMSE moyen des températures de surface du lac de tous les forçages avec le RMSE d'un cas spécifique, nous pouvons observer une forte dispersion sur le nuage de points. L'augmentation de la résolution du forçage a un impact, le RMSE diminue de  $0.5^{\circ C}$  quand le forçage appliqué est à une résolution spatiale de 0.25 degrés par rapport à 0.5 degrés. De même sur la période de gel, il peut y avoir un écart de 2 mois entre les différentes simulations comme le montre les simulations du lac Baikal. Les erreurs médianes sur la période de gel varient entre 30 et 40 jours pour les forçages à 0.5 degrés et en dessous de 25 jours pour les forçages à 0.25 degrés.

Les erreurs sont assez peu influencées par la stratégie de paramétrisation de la profondeur. Seule la date de début du gel du lac est impactée par cette stratégie. Mais le passage en prenant en compte trois types de lac dégrade la prévision.

Plusieurs processus non modélisés comme la variation temporelle de la profondeur, la salinisation du lac ou encore la paramétrisation du coefficient d'extinction peuvent augmenter les erreurs du modèle. Ainsi, il sera intéressant de se focaliser sur ces aspects pour améliorer le couplage.

### 5.4 . Article 2

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

Bernus Anthony and Otlé Catherine

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

**Correspondence:** Anthony Bernus (anthony.bernus@lscce.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. The different categories 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  $0.5^\circ$  or  $0.25^\circ$  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 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 range between  $3.2^\circ\text{C}$  and  $2.7^\circ\text{C}$  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  $0.1^\circ\text{C}$  for the four aggregation scenarios tested.

The simulation of ice phenology presented systematic errors whatever the forcing used and the depth parameterization. Freezing onset was shown to be the less sensitive to forcing and depth parameterization with the median of the errors ranging between 10 and 14 days. Larger errors were observed on the simulation of the end of the freezing period significantly influenced by the atmospheric forcing used. Such errors already highlighted in previous works, could be the result of deficiencies in the modeling of snow/ice parameterization processes. 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

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 to the increased grid resolutions and computer advances.

Lake distribution is spatially unequal all over 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 (Messenger et al., 2016). Their role on the physical properties of the low atmosphere and the importance to model these 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  $2^{\circ}C$  to  $3^{\circ}C$  and that the latent heat fluxes increased by 10 to  $45 W/m^2$  depending on the region.

Lakes play also a buffer role in the hydrological cycle impacting substantially the river regimes. As an example, in Alaska, 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 (Cole et al., 2007). Another important GHG is the methane from methanogenesis by micro organisms in anaerobic condition, processes which are presently poorly understood but could lead to lake emissions between 8 and  $48 Tg C - CH_4 y^{-1}$ , 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-anthropogenic methane emissions are coming from lakes. Recently, Beaulieu et al. (2019) gave a higher estimation of lake methane emissions with an amount of  $112 Tg C - CH_4 y^{-1}$ .

Given the role of lakes on the global carbon and water cycles, the large uncertainties related to GHG budgets (Saunio 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. This implementation requires as a first step to solve the surface energy balance to be able to predict the time evolution of the surface temperature and fluxes exchanged with the atmosphere. Among 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 with a little number of free parameters. In this category, the FLake model (Mironov, 2008) and the Hostetler 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 developed by MacKay (2012) resolving the turbulence in the mixing layer with a bulk approach may also be mentioned. These three models have shown good performances when compared at site level in various intercomparison exercises such as LakeMIP (Stepanenko et al., 2010, 2013).

As such, FLake has been implemented in numerous climate and Numerical Weather Prediction models (NWP). For example, it has been coupled to the German regional atmospheric model (COSMO) and its TERRA land component (Mironov et al., 2010), the European ECMWF atmospheric forecasting model and its HTESSEL land surface component (Balsamo et al., 2012; Dutra et al., 2010), the UK Met Office Unified Model and its JULES Land simulator (Rooney and Bornemann, 2013), and the French CNRM-CM5 ESM and its SURFEX interface (Le Moigne et al., 2016). The coupling of FLake with all these LSMs showed strong benefits in the climate and weather simulations. In addition, the LISS model based on the Hostetler model is used in the CLM land component of the CESM climatic model developed by the NCAR institute (Oleson et al., 2013) and in the ELM land model based on the same core of CESM as BCC-AVIM, (Golaz et al., 2019) and NorESM2 (Seland et al., 2020). The LISS model is also implemented in LM4 which is the land component of the GFDL 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 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 IPSL ESM and more precisely in its land surface component ORCHIDEE. Contrary to Krinner (2003) who directly introduced the Hostetler model in the atmospheric model component (LMDZ) of the ESM 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 have been done to assess the lakes features 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 the 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

90 The ESA Land Cover CCI product (v2.0.7, Bontemps et al. (2015), ESA (2017) ) developed in the framework of the Cli-  
mate Change Initiative program of the European Space Agency (ESA) was used to map the natural vegetation and derive  
the ORCHIDEE Plant Functional Types (PFTs). 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 OR-  
CHIDEE PFTs is fully presented in Poulter et al. (2015) and Hartley et al. (2017), and detailed in the dedicated website  
95 <https://orchidas.lscce.ipsl.fr/dev/lccci>. In this work, we used the PFT maps generated at the resolution of  $0.25^\circ$  and  $0.5^\circ$  for the  
standard version of ORCHIDEE.

#### 2.1.2 Lake area and depth

The HydroLAKES database (Messenger 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  $0,1 \text{ km}^2$ . A mean depth is provided for each water  
100 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 insure  
the coherence for the following developments when lakes will be connected to the river routing. The data can be downloaded  
105 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 simula-  
tions. This dataset has been developed thanks to a project funded by the Natural Environment Research Council (NERC) in the  
continuity of the ARC-Lake database (MacCallum and Merchant, 2012) funded by the European Space Agency. It provides  
110 daily observations of Lake Surface Water Temperature (LSWT) for about 1000 lakes distributed all other the world, with uncer-  
tainties and quality level. The observations cover the period 1995-2016 with a spatial resolution of  $0.05^\circ$  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  
115 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 use the 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) is less than 1 K. The comparison against local measurements show that errors are slightly larger for day time measurements (0.7 K) compared to night time (0.5 K) with a small negative bias ( $-0.2^{\circ}\text{C}$  during day time and  $-0.1^{\circ}\text{C}$  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 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. The data are available for some lakes back to the year 1850 (Benson, 2002). The data can be downloaded at <https://doi.org/10.7265/N5W66HP8>.

#### 2.1.5 Atmospheric forcings

Different atmospheric reanalysis are used 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. In this work, we used five different datasets available at three different spatial resolutions: 31km,  $0.25^{\circ}$  and  $0.5^{\circ}$  and three different temporal resolutions, hourly, 3-hourly and 6-hourly.

Then, for the simulations performed at  $0.5^{\circ}$  resolution, we used: a) WFDEI-CRU, a 3-hourly  $0.5^{\circ}$  global forcing product 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  $0.5^{\circ}$  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  $0.5^{\circ}$  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 product at a resolution of 31 km, and E2OFD (Beck et al., 2017), a 3-hourly  $0.25^{\circ}$  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: <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 times, follows the typical daily evolution according to the solar zenithal angle. It should be noted that no extra atmospheric fields are needed to solve the energy balance at the lake – atmosphere interface, same as what is done for the vegetation part.



## 145 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. 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 versions of ORCHIDEE, urban areas and water surfaces are not represented explicitly and are characterized by a fraction of 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) and more recently in Harper et al., in preparation. 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 exercise, and contributed to the IPCC 6<sup>th</sup> 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 developed for NWP 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). 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 free water bottom temperature, the thermocline shape factor, the ice layer top temperature and ice thickness.

180 If the snow module is activated, the snow temperature at the atmosphere interface and the snow depth are calculated, based on the bulk resolution of the heat budgets of the snow layer. It should be noted that the model resolves only the energy budget equations and that the water balance is not solved. That means that the water volume is kept constant in time. The depth is therefore input parameters of the model. In addition, two important radiative parameters (surface albedo and light extinction coefficient) and the surface fetch involved in the calculation of the surface fluxes need to be prescribed to run the model.

185 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 with a constant temperature around  $4^{\circ}C$ . 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

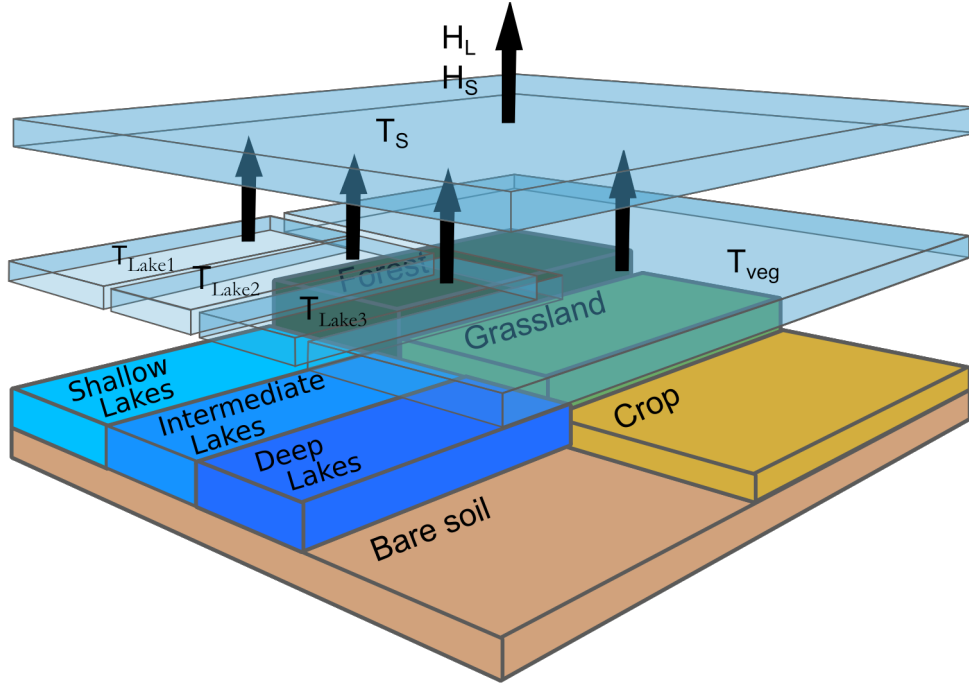
#### 3.1 Integrating FLake into ORCHIDEE

190 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, 195 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.

#### 200 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 205 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.



**Figure 1.** Lakes tiles representation in ORCHIDEE land surface model

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

$\Phi_{\Theta}$  represents the normalised temperature depending of the normalised depth  $\zeta$  as follows:

$$\Phi_{\Theta} = \frac{T_s - T(z)}{T_s - T_b}$$

$$\zeta = \frac{z}{D} \quad (2)$$

210 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  $z$ .

Mironov (2008) defined a shape factor  $C_{\theta}$  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  $C_{\theta}^{\max}$ .

$$215 \quad \frac{dC_\theta}{dt} = \text{sign}(dh/dt) \frac{C_\theta^{\max} - C_\theta^{\min}}{t_{rc}}, \quad C_\theta^{\min} \leq C_\theta \leq C_\theta^{\max} \quad (3)$$

Some issues were reported by Salgado and Le Moigne (2010) who found unrealistic behavior of the temperature profile in 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  $0.01 \text{ h}^{-1}$  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 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 assumption of the temperature profiles. At the snow-ice interface, Equation 4 is used to calculate the snow temperature and derives from the condition of continuity of the heat flux.

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

$R_h$  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 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 snow module to prevent the numerical instabilities with the advantage of preserving the water budget. Various tests were performed to adjust the time split factor which has been finally set to a value of 50. Therefore, when the snow thickness is less than 3 cm, the time step of the resolution of the snow module is decreased by a factor of 50. 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 the Bear Lake 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 a constant value of  $320 \text{ kg} \cdot \text{m}^{-3}$ . Their results show the benefit of parameterizations accounting for snow aging, but given that the use of on 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  $k_s$  was given by Equation 5 with  $k_{s_{max}}$  and  $k_{s_{min}}$  the maximum  
 245 and minimum values of the snow heat conductivity and  $h_s$  the snow depth. It was replaced by Equation 6, where  $k_s$  is set to  
 0.14  $W m^{-1} K^{-1}$ ,  $k_i$  is the heat conductivity of the lake ice (2.29  $W m^{-1} K^{-1}$ ) with  $c$  an empirical constant (5  $m^{-1}$ ).

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

$$k_{seff} = k_s + (k_i - k_s) \cdot e^{-c \cdot h_s} \quad (6)$$

Following Semmler et al. (2012), Pietikäinen et al. (2018) revised also the calibration of the ice and snow albedos and  
 250 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.

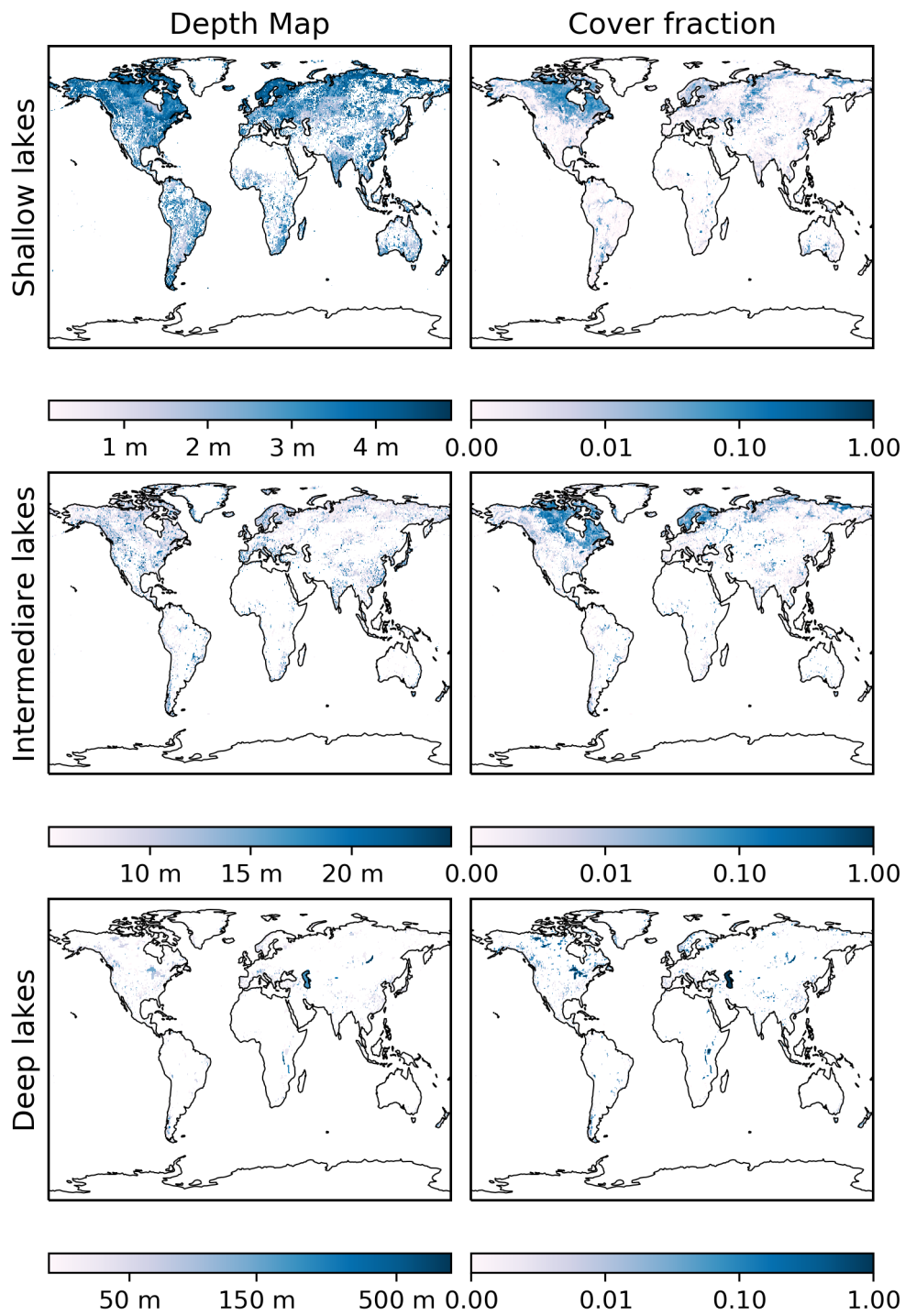
### 3.3 Model settings

ORCHIDEE requires the description of the land surface variability in each model grid cell to simulate the energy and biogeo-  
 chemical processes. As already noted, the standard version uses PFT maps derived from the CCI Land cover product. Since  
 255 the standard PFT maps do not consider water surfaces, the permanent water bodies 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  
 260 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.

#### 265 3.3.1 Lake surface area and depth

The HydroLAKES database (Messenger 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  
 270 depths and aggregated in a single class. The median values of the corresponding depths are shown on the same figure. The  
 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 around 20° – 30° North and South, in coherence



**Figure 2.** Lake fractions (right column) and lake depth (left column) averaged for each lake categories at 0.5°

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, the Victoria Lake in South Africa, the Malawi Lake in South America or the Caspian Sea and Baikal Lake in Eurasia.

### 3.3.2 Other lake parameters

FLake requires the prescription of other parameters like the radiative parameters of water (surface albedo and extinction coefficient) 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  $1 \text{ m}^{-1}$  was set for water. The fetch parameter which appeared less influent in the sensitivity analysis performed by Bernus et al. (2021), has been parameterized according to the surface area of the lake fraction within the grid cell. Considering a circle pattern with the same area than the considered lake, the fetch value has been set equal to the diameter of the circle.

### 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 which are considering as bare soil 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 2000-2016 at a time step of 30 min. A spin up procedure of 20 years in the previous years was applied to bring the prognostic variables to equilibrium. 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 the related uncertainties. The performance metrics, i.e. the RMSE and its components, used in the evaluation are presented in the Appendix A.

**Table 1.** List of the ten ORCHIDEE-FLake simulations performed at global scale, differing from the atmospheric forcing used and the number of lakes considered

ID	Meteorological Forcing (spatial and temporal res.)	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

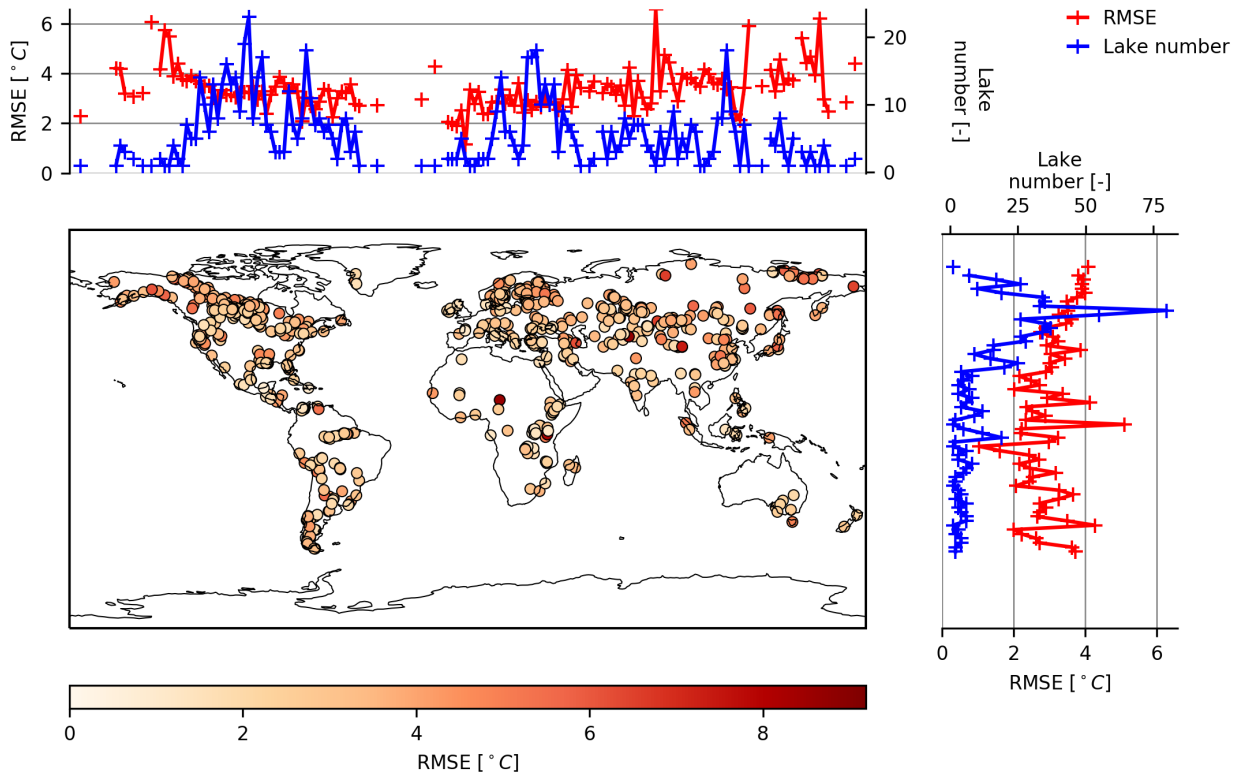
## 4 Results

### 4.1 Evaluation at global scale

#### 4.1.1 LWST

The simulated LWST for CJR3med have been compared to the observations provided by GloboLakes for the the 979 lakes  
305 matching the 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 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. This sampling is representative of the area lake distribution plotted in  
310 Figure 2. The RMSE distribution calculated with GloboLakes is therefore a good representation of the ORCHIDEE-FLake  
model error. Nevertheless, GloboLakes does not provide LWST for frozen lakes, thus LWSTs below  $0^{\circ}C$  are not used in the  
RMSE calculations. The results show that most of the RMSEs are ranging between 2 and 4 K. Higher errors can be seen for a  
few lakes with RMSE values up to 8 K. The latitudinal representation highlights a positive trend of the RMSE, when moving  
from the temperate to the boreal regions. In the tropics, below  $20^{\circ}N$ , some large RMSE peak values are observed, which are  
315 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.





**Figure 3.** LWST errors (RMSEs) simulated by ORCHIDEE-FLake in the CJR3med configuration run at  $0.5^\circ$  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.

#### 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. In Table 2, the mean, quartiles and median of the differences between simulated and observed variables are given. The evaluation is shown for three output variables of the CJR3med simulation, which are the start and end of the freezing period (SOF and EOF respectively), and its duration, all in days unit. The results show that the SOF errors are all negative and the EOF 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 SOF is about -11 days (median value), 21 days for EOF, leading to an ice duration overestimated by 32 days in average. The RMSE values are around 24 days for both SOF and EOF and reach 40 days for the ice duration error. The interquartile range is larger for SOF (19 days) compared to EOF (10 days) highlighting the fact that the SOF errors are more scattered. This also explains why the differences between the median and mean values are larger for SOF compared to EOF.

**Table 2.** Statistical metrics (simulations - observations) of the lake ice phenology (CJR3med configuration)

Statistical metrics	Start of Freezing (days)	End of Freezing (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

330 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 LWSTs 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). The LWST time series (simulated and observed over the year 2012) are plotted in Figure 4 along with  
335 the RMSE 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 LWST, since GloboLakes observations do not consider frozen conditions.

Figure 4 shows that the seasonal variations of LWST are generally well represented for all the lakes. The RMSE values are  
340 ranging between 2 and 4 °C but can show large variations according to the atmospheric forcing. The dispersion is the largest for the two tropical lakes (Victoria and Titicaca) for which LWST biases can reach 6 K. Biases explain a large part of the RMSE error for the other lakes, especially for Erie and Superior lakes. These results 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.

The LWST errors of the tropical lakes are principally explained by biases compared to the other climates because of the low  
345 seasonal temperature amplitude in these regions. The values of the RMSEs range between 1 and 4 °C among the lakes. For a given lake, large variations of the RMSEs are observed between the 5 different forcings. For example, the RMSE of Lake Victoria varies between 1 to 4 °C depending on the meteorological forcing. The bias SB is the component which could vary the most between each time series (up to 4 °C) for the tropical lakes (Victoria and Titicaca). The second component which is significant is the standard deviation LCS, which could reach 1 °C (particularly for the lakes under cold climates). The last  
350 component SDSD which allows to diagnose how the model reproduces the LWST amplitude, is always small, less than 0.5 °C.

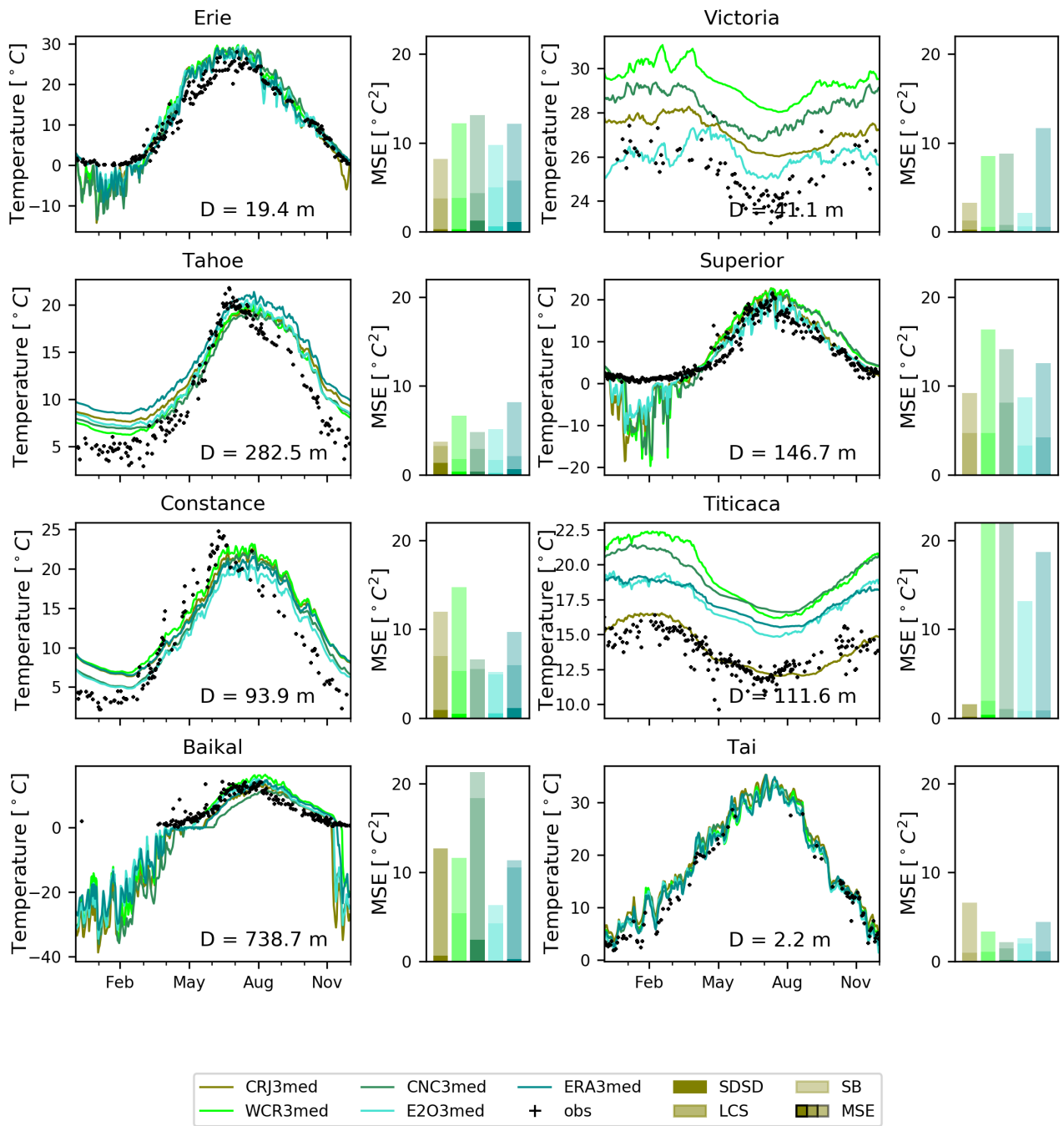
For the temperate lakes, the annual surface temperature amplitude is larger than for the tropical ones, between 15 and 20 °C. Tai Lake is well represented with RMSEs below 2 °C for most of the forcings. For the two other temperate lakes (Tahoe and Constance), the model fails to simulate the annual amplitude which is underestimated and the cooling at the end of summer is delayed whatever the forcing used. Such errors could result from parameterization errors, such as depth errors which are known to influence greatly the seasonal LWST pattern. Even if the freezing periods are not accounted in the RMSE calculations, the time series highlight the large impact of the atmospheric forcing on the timing and amplitude of the ice cover. For example, if we consider the Erie Lake which did not freeze in the observations in year 2012, the winter temperatures are very well modeled with all the forcings except when CRUNCEP is used. This failure has been explained by the cold bias of the air temperature identified in CRUNCEP over this part of Canada when compared to the other forcing datasets (not shown here).

#### 360 4.2.2 Global scale analysis

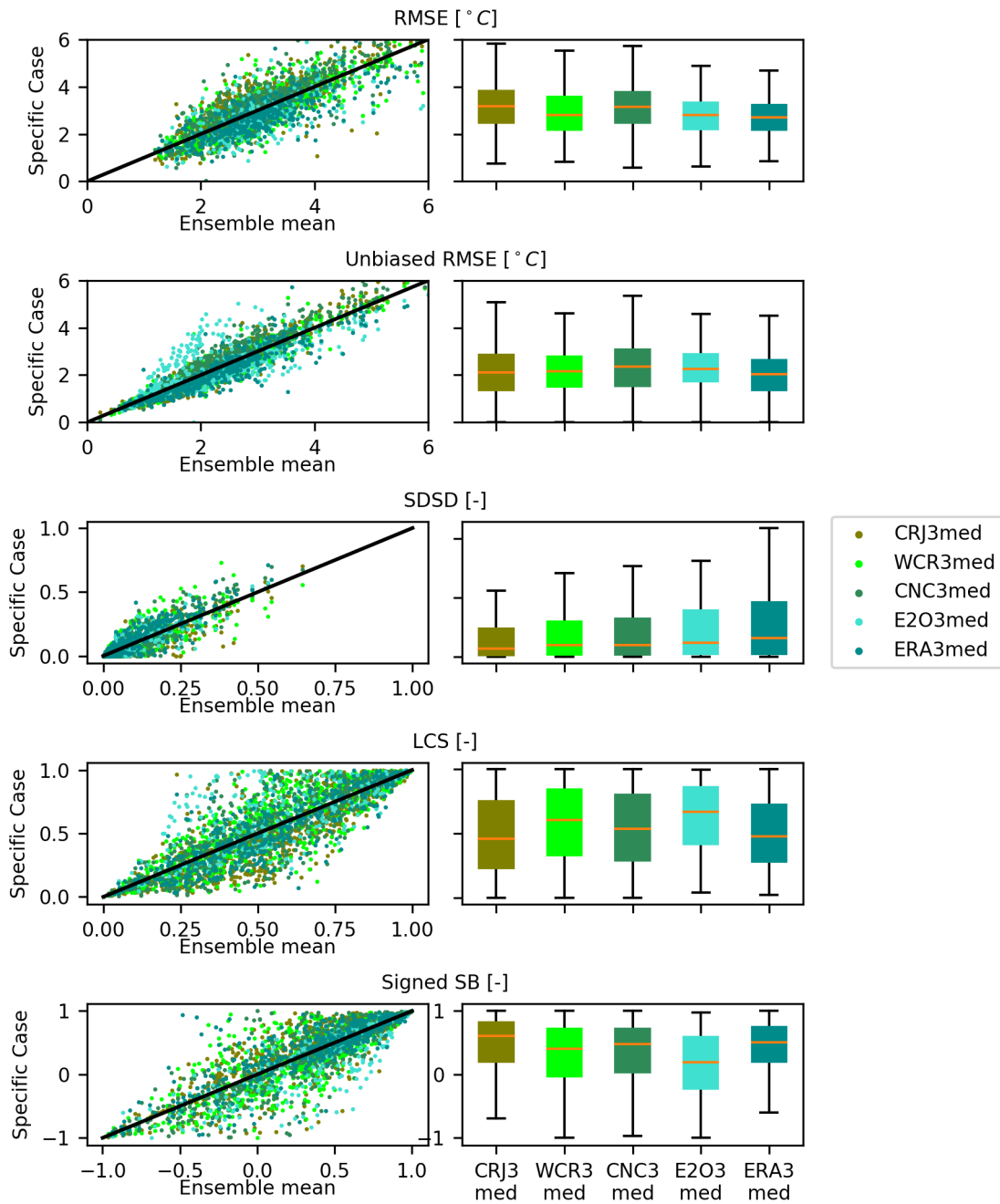
Figure 5 presents the performance metrics calculated on all the lakes observed in GloboLakes that were represented in ORCHIDEE in the corresponding depth class tile (i.e., 979 lakes all other the world). The statistics were calculated for the seventeen years of the whole simulation period (2000-2016). The metrics shown here are the RMSE, the unbiased RMSE, and the three components of the total RMSE, i.e., SB, SDSD and LCS. The values obtained for each lake and for each simulation are compared to the ensemble mean in the scatterplots. Therefore, these plots inform on the sensibility of the surface temperature to the atmospheric forcing. In addition, the boxplots show the metrics calculated for each simulation forced by the five atmospheric forcings used.

The results show that the simulations forced by the meteorological reanalysis with the lower resolution (i.e., 0.5°) lead to the worse 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 K. An important part of the errors is therefore explained by the resolution and accuracy of the meteorological forcings. Thus, the model errors will be more complicated to evaluate.

However, the decomposition of the RMSE is quite similar among the forcings on average. The scatterplots on the left panel compare the mean RMSE for the different forcings. The high dispersion on the scatterplot shows that the modeled surface temperature is sensitive to the uncertainty of the atmospheric forcing. The LCS and SB components explain about 50% of the RMSE whatever the forcing used, the SDSD component being minor. Furthermore, LCS and SB explain at least 50% of the RMSE for more than 50% of the observed lakes when SDSD explains less than 20% of the RMSE for more than 75% of the observed lakes. Thus, the LWST 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, the signed bias is plotted in the scatter plot whereas in the whisker plot the squared bias (SB) normalized as the other components is plotted. Then, it is clear that for all the forcings, the bias is mostly positive, meaning that the model generally overestimates the surface temperature. Compared to the total RMSE that can reach up to 10 K, the unbiased RMSEs are reduced and do not exceed 6 K. The dispersion is also less important than the one observed with the standard RMSE. Indeed, the standard deviation of the



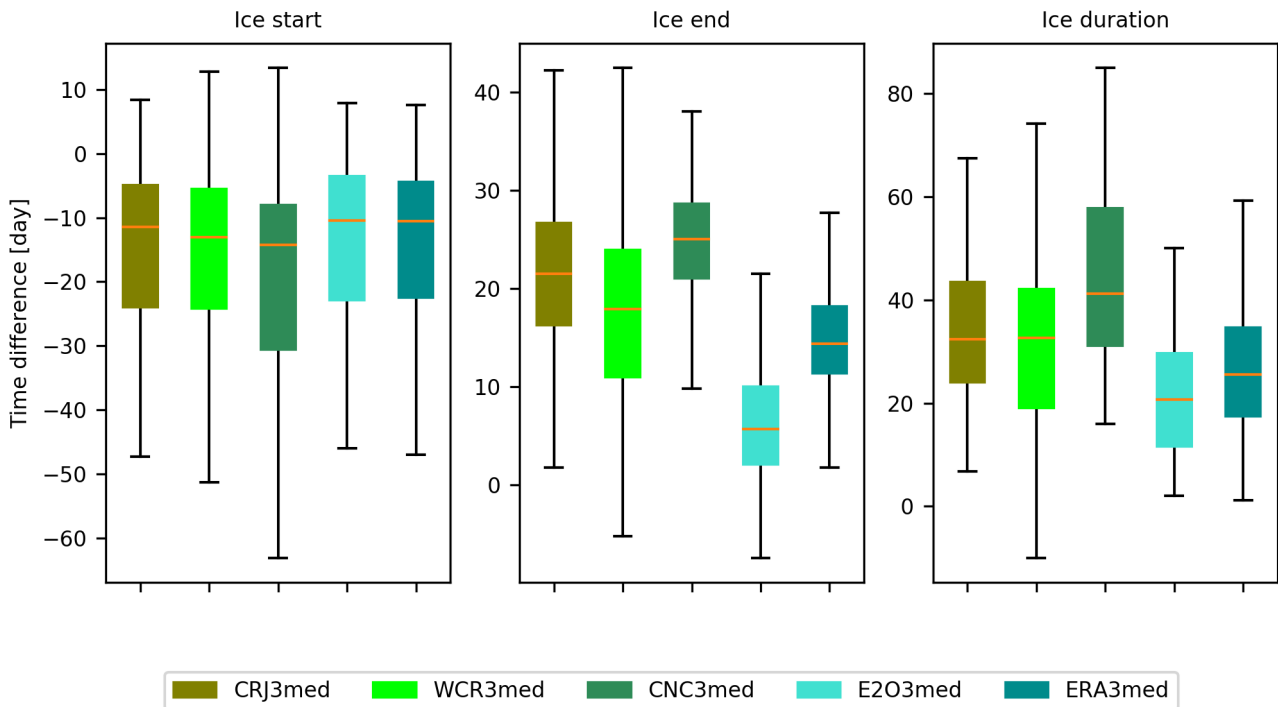
**Figure 4.** Time series of the daily mean lake surface temperature simulated for eight lakes with five different meteorological forcings compared to GloboLakes observations for the year 2012. The barplots represent the decomposition of the RMSEs in terms of bias, lack of correlation and difference of standard deviations between the model and observations.



**Figure 5.** LWST errors obtained on the 979 lakes observed in GloboLakes for the five simulations CRJRA3med, WCR3med, CNC3med, E2O3med and ERA3med. RMSE, unbiased RMSE, LCS, SSDS and sign SB are represented as scatterplots (left panel) in comparison to the ensemble mean and in boxplots (right panel) where the median is represented by the orange line, and the quartiles by the edge of the color box

values to the identity line ( $\sqrt{\frac{1}{N} \sum (RMSE_i - RMSE_{Mean})^2}$ ) is  $0.47 \text{ } ^\circ\text{C}$  on average for the RMSEs and  $0.20 \text{ } ^\circ\text{C}$  on average  
 385 for the unbiased RMSEs. Therefore, the model biases explained a large part of the LWST errors.

The sensitivity of the simulated ice phenology (SOF, EOF and ice duration) 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 (SOF) than on the prediction of the end of the freezing period (EOF), 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  
 390 with the higher resolution forcings (E2OFD and ERA5) with a significant advantage of E2OFD especially for the simulation of EOF. The SOF, EOF and duration median errors are respectively equal to -6, 11 and 20 days with E2OFD whereas they were up to -14, 25 and 41 days with the CRUNCEP forcing.

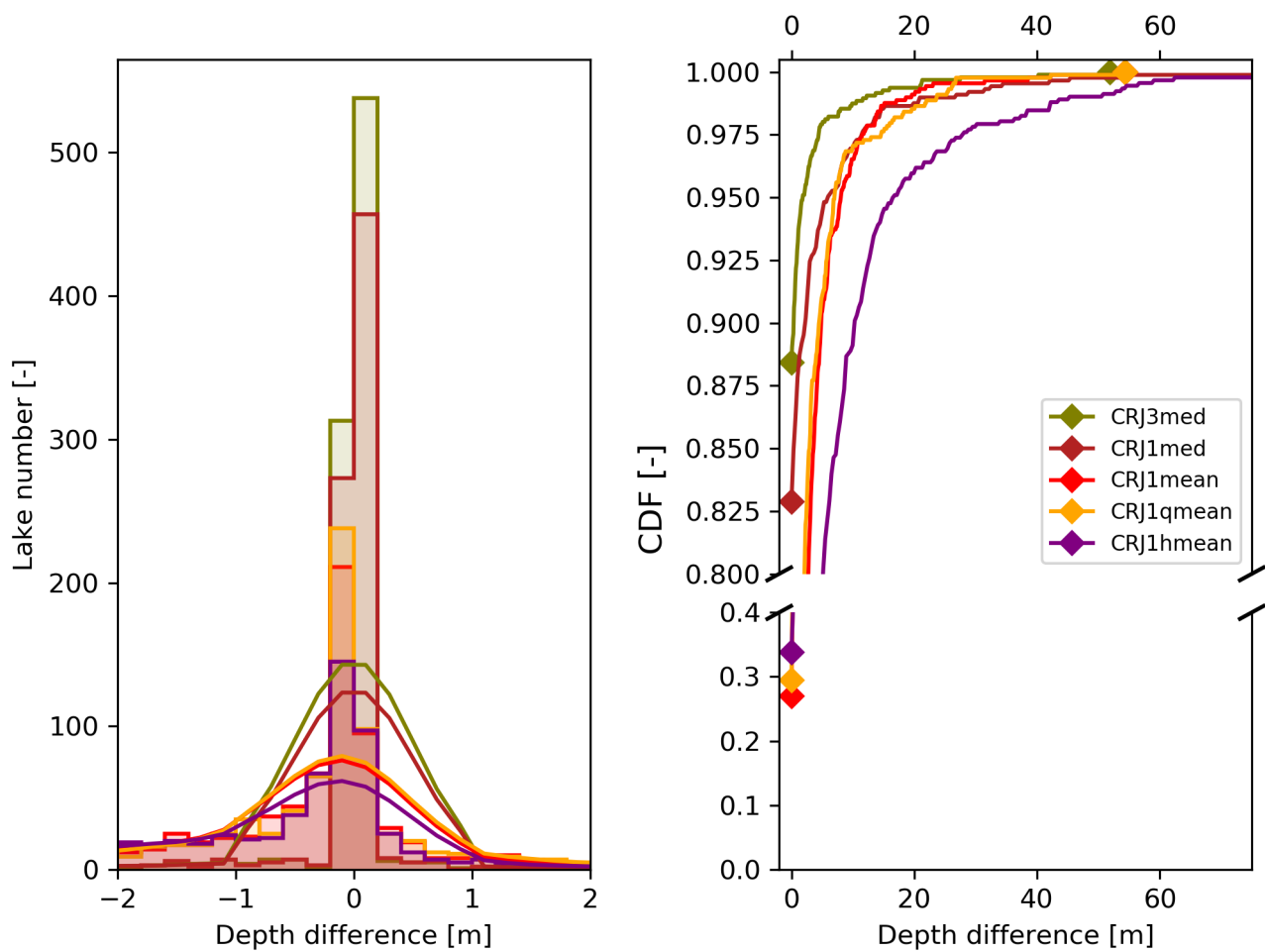


**Figure 6.** Boxplot of the RMSE errors on the start, end and duration of the ice period for the five simulations CJR3med, WCR3med, CNC3med, E2O3med and ERA3med.

### 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 observed  
 395 lakes and the depth prescribed in the corresponding ORCHIDEE tile at a grid resolution of  $0.5^\circ$ . A preliminary step to assess this error is to assign the lake from HydroLAKES to the lake from GloboLakes. We succeed 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 are in agreement over a single lake, the depth referenced in HydroLAKES was assigned as the observed depth. But when several water bodies in HydroLAKES correspond to a single observed lake, we assign the area weighed mean depth to assess the depth of the observed lake. Figure 7 presents the distribution of the errors between the observed depth as defined previously and the depth prescribed in the 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 probability density functions (PDF) and the cumulative distribution functions (CDF) are both plotted.



**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.

405 The configuration with 3 tiles shows the lowest errors on the depth parameterization, and the median appears to be the best way of depth aggregation within a grid cell: about 90% of the observed lakes have an absolute error 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 lakes have an absolute error less than 4,4 m the quadratic mean and than 4.9 m with mean depth estimation approach and the harmonic mean is the worst approach since only 90% of the observed lakes have an absolute error less than 10.2 m.

410 In Figure 8, the total RMSEs, the unbiased RMSEs, SDS, LCS and the signed bias are plotted as scatterplots against the ensemble mean and as box plots. Compared to the previous sensitivity analysis to the atmospheric forcing, the errors are slightly lower as well as the dispersion among the simulations. The RMSE values are generally lower than  $6^{\circ}\text{C}$  with a median value between  $3.1^{\circ}\text{C}$  and  $3.2\text{ K}$ . The components present similar order of magnitude among the simulations and the scatterplots show a low dispersion around the identity line for each component. This result was expected since the impact of the lake depth  
415 parameter in FLake, impacts mostly the pattern of the seasonal cycle of LWST and to a lesser extent, its amplitude.

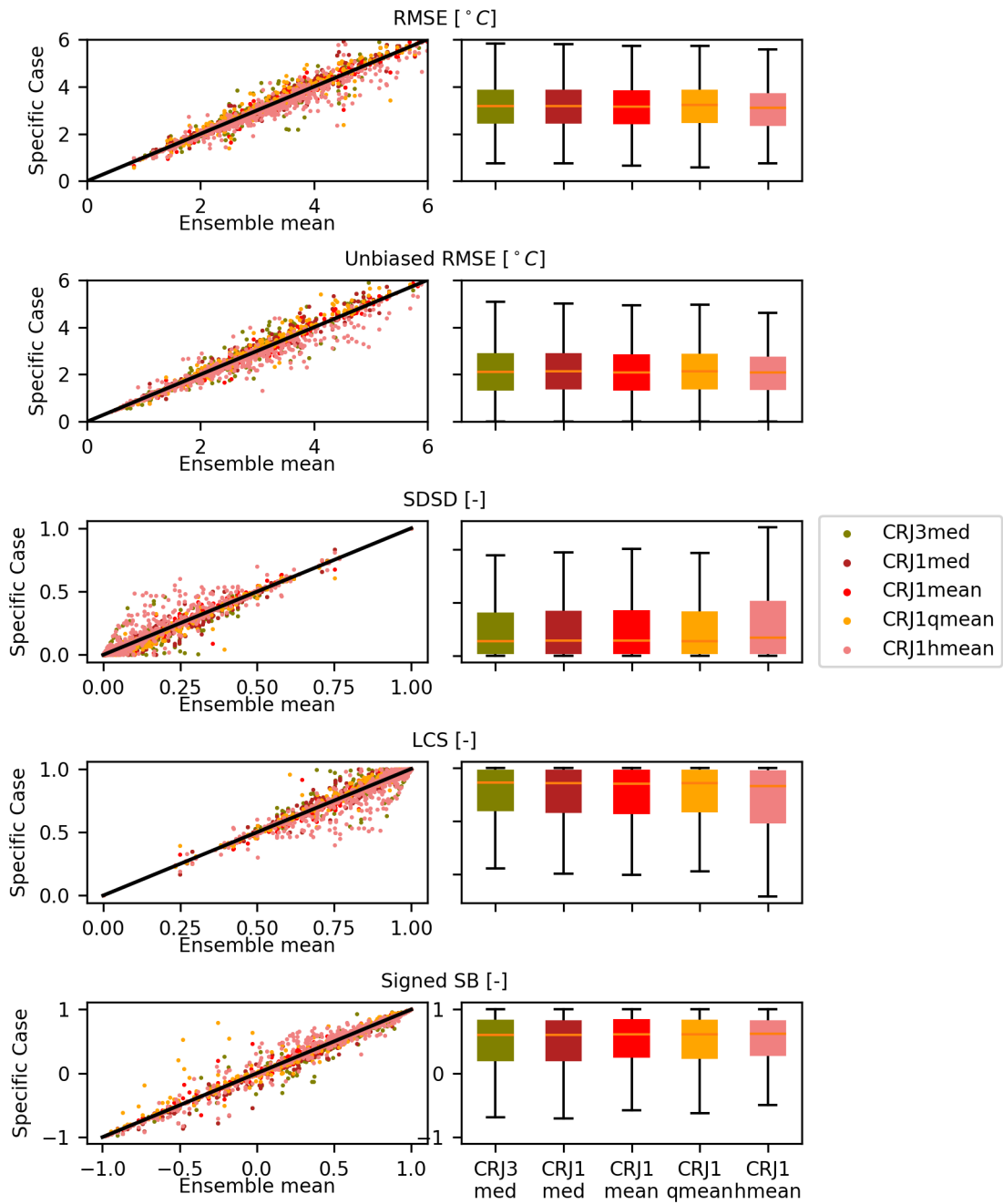
The ice phenology sensitivity to the number of lake tiles and depth parameterization strategy is presented in Figure 9. Here also, the results show less impact compared to the sensitivity obtained with the meteorological forcing. In contrast, it appears that the SOF errors are reduced by half in the 1-tile configurations compared to the 3-tile one. In the four 1-tile simulations, SOF errors are around 5 days whereas it is equal to 10 days in the CJR3med simulation. However, the variance is quite  
420 similar, the 3-tile simulation shows a clear bias not observed in the 1-tile simulations. EOF and ice duration errors are similar among the five simulations, equal to about 22 days and 27 days respectively, comparable to what is obtained in the CJR3med simulation. The improvement observed in the simulation of SOF 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 SOF time: LWST cooling at the end of summer is too rapid and the lake freezing is earlier compared to  
425 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 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 compensate the too early cooling. The impact of this error is less visible on the EOF because the depth parameter is less influential during the ice melting  
430 process (Bernus et al., 2021).

#### 4.4 Energy balance impacts of lakes in ORCHIDEE

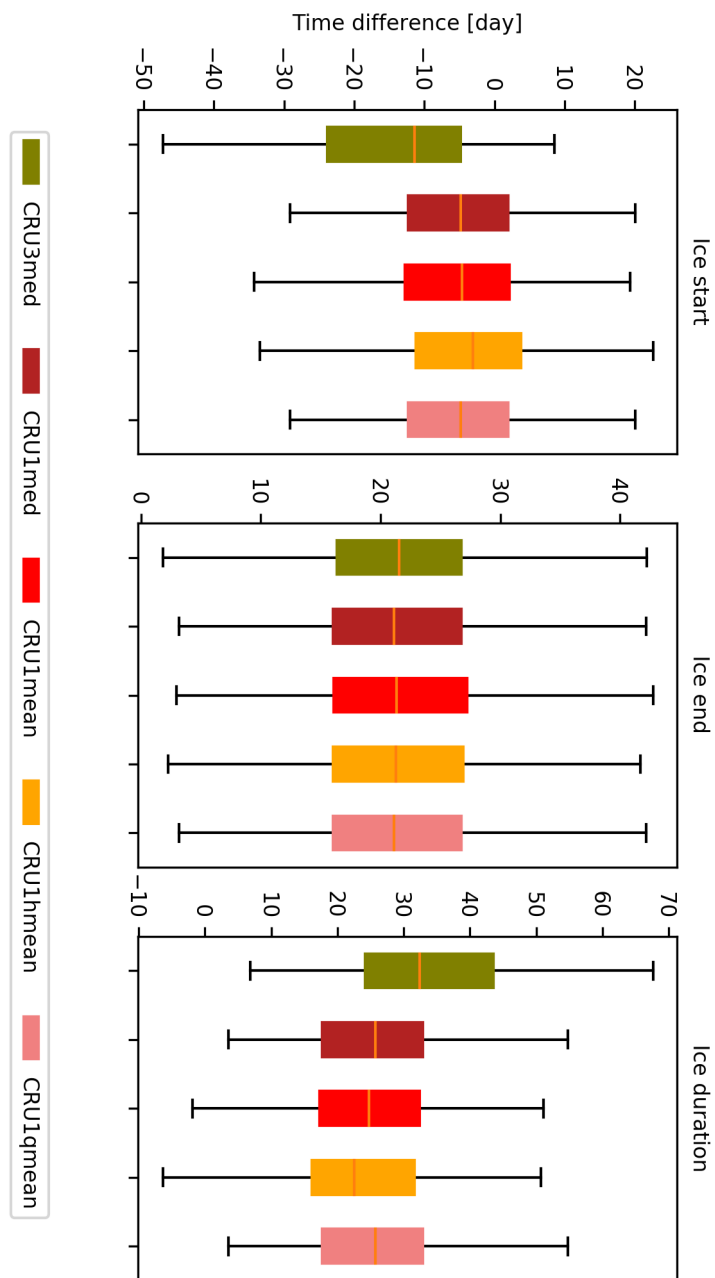
To analyse the impact of representing lakes on the global energy and water budgets, the two simulations CJR0 et CJR3med were compared at the global scale. The differences (CJR3med-CJR0) in terms of surface temperature, latent and heat fluxes averaged over the grid cell, were calculated for the four seasons of the year and plotted in Figure 10. Winter period is defined  
435 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 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

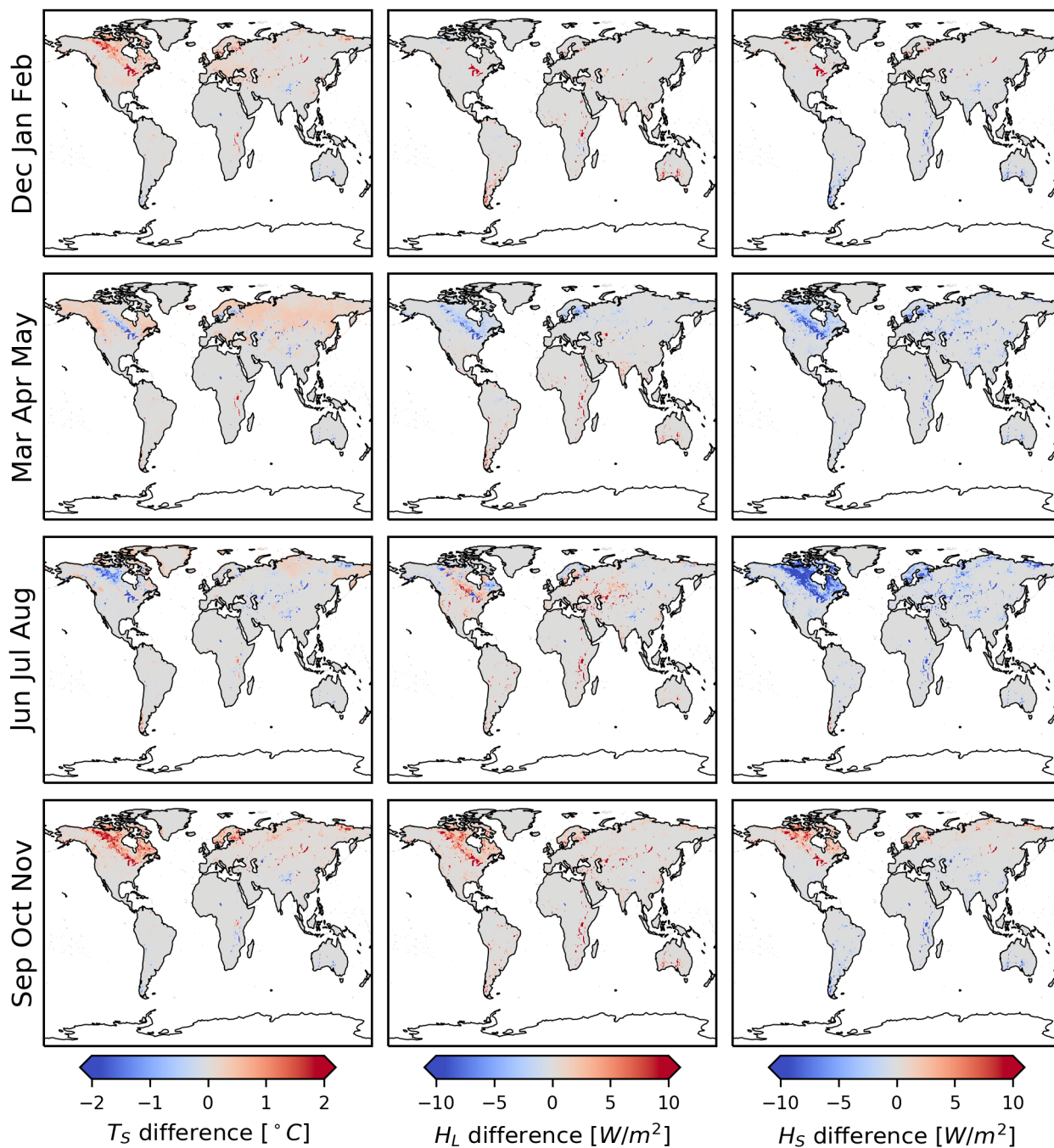




**Figure 8.** LWST errors obtained on the 979 lakes observed in GloboLakes for the five simulations CJR3med, CJR1med, CJR1mean, CJR1qmean and CJR1hmean. RMSE, unbiased RMSE, LCS, SDSD and SB are represented as scatterplots (left panel) in comparison to the ensemble mean and in boxplots (right panel)



**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



**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)

warmer for all kind of lakes. For the deeper lakes, this difference can reach 10 K. It is the case for the Lake Baikal whereas for  
440 the Michigan and Superior lakes, the difference is between 5 and 7 K. In the regions covered by smaller and shallower lakes  
in Scandinavia and Canada, this difference remains below  $2^{\circ}\text{C}$  due to the lower lake fraction and smaller smoothing effect of  
shallower lakes. In spring, the deep lakes as the great American lakes or the Lake Baikal lake are cooler than their environment  
and explain the differences of temperature between the two simulations by  $5\text{-}6^{\circ}\text{C}$  and  $3\text{-}4^{\circ}\text{C}$  respectively. However, the  
445 regions where the lakes are shallower remain warmer in the lake simulation because of their lower thermal inertia (but only  
by 0.5 K). 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 October and April, the simulation with lakes is warmer at the end of the humid season by 1 or  $2^{\circ}\text{C}$  (in Africa  
during the spring season) and cooler during the end of the dry season by 0 or  $3^{\circ}\text{C}$  (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.  
450 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 grid cell completely covered by water, i.e., lake fraction is equal to 1, the evaporation difference can reach 0.31 mm/d. In  
tropical areas, during the summer period and for latitudes below  $30^{\circ}\text{N}$ , 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.

455 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  
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 Baikal Lake is between  
25 and  $50\text{ W/m}^2$  less than the bare soil evaporation and it is about the same for the Lake Superior (between 20 and  $40\text{ W/m}^2$ ).  
460 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.

## 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  
465 to evaluate the model sensitivity to the meteorological forcing and to the lake tile depth parameterization. The evaluation of  
LWST and ice phenology against the GloboLakes dataset and the Global Lake and River Ice Phenology Database, showed that  
the lake surface temperature is quite well simulated in this new ORCHIDEE version. Compared to the observations, LWST  
errors range between 2 and 4 K, even if in some few cases, the errors can reach 8 K. These results are comparable to what  
was obtained by previous works, such as Dutra et al. (2010) who obtained the lake RMSE median around  $3.75^{\circ}\text{C}$  in the high  
470 latitudes ( $> 40^{\circ}$ ) and around  $2.25^{\circ}\text{C}$  for low latitudes ( $< 40^{\circ}$ ). 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, compared to previous studies, the results obtained are quite similar. As an example, Dutra et al. (2010) obtained a SOF error of 30 days (in advance) for 75% of lakes after optimisation of the depth parameter against LWST 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  $0.25 \text{ mm/d}$ ) and smoother LWST 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 the same 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 LWST 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 K. For those lakes, we observed strong biases in the atmospheric temperatures (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 LWST 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 LWST 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 ice phenology. The large sensitivity of the SOF/EOF 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 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 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 LWST errors were smaller when the two best resolution forcings are used (E2OFD and ERA5). Increasing the spatial resolution improved LWST RMSEs by 0.5 K. 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 LWST 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 (SOF). As already said, lake depth is determinant on the lake freezing process but less on the melting one which are more driven by the atmospheric conditions and the amount of snow than directly on the lake features.

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 LWST and ice phenology. Lakes 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. For example, the Lake Chad in HydroLAKES has a surface area doubled compared to the current value and the average depth is probably erroneous as well. If we consider the ice observations, SOF time corresponds to the start of freezing in the observations 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 totally frozen in a time step and so abruptly will vary the related radiative parameters. For big lakes with large surface areas, this scaling error may be important. Similar conclusions were drawn by Pietikäinen et al. (2018) who identified two factors explaining the delay in their simulated EOF and the advance of SOF with the coupled FLake-REMO model. Firstly, the lakes are freezing too early and 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. 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. Secondly, 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.

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, 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 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 observe for the Urmia lake in Azerbaijan. The consideration of salinity in the water freezing point could improve the results, especially the cryospheric processes.

### 540 **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  
545 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 come 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  
550 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  
555 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 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  
560 environmental changes.

*Code availability.* Code for this project is available at [https://forge.ipsl.jussieu.fr/orchidee/wiki/GroupActivities/CodeAvailabilityPublication/ORCHIDEE\\_FLAKE\\_gmd\\_2021](https://forge.ipsl.jussieu.fr/orchidee/wiki/GroupActivities/CodeAvailabilityPublication/ORCHIDEE_FLAKE_gmd_2021).

### **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  
565 the Root Mean Square Error (RMSE) which is widely used to measure the performance of a model and is defined by Equation A1.

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

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). 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 add the signed squared bias (SSB)

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

with

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

$$LCS = \frac{\sqrt{\frac{1}{N} \sum (\overline{T_{sim}^F} - T_{obs})^2}}{\frac{1}{N_F} \sum (\sqrt{\frac{1}{N} \sum (T_{sim}^F - \overline{T_{sim}^F})^2})} \quad (A4)$$

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 A5. We also used this metric in our study.

$$RMSE_{unbiased} = \sqrt{\frac{1}{N} \sum (T_{sim} - T_{obs} - \text{sign}(\overline{T_{sim}} - \overline{T_{obs}}) \sqrt{SB})^2} \quad (A5)$$

580 *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'Énergie Atomique et aux énergies alternatives) is also acknowledged 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.

## 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.
- 590 Balsamo, G., Salgado, R., Dutra, E., Bousseta, 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.
- 595 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, 2017.
- 600 <https://doi.org/10.5194/hess-21-589-2017>, 2017.
- 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.
- 605 Bernus, A., Otle, C., and Raoult, N.: Variance based sensitivity analysis of FLake lake model for global land surface modeling, *Journal of 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](https://doi.org/10.1175/1520-0442(1995)008<2691:SOAGST>2.0.CO;2), 1995.
- 610 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.
- 615 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., Levvasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luysaert, 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.
- 625 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.
- 630 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.
- 635 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.
- 640 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](https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf), 2017.
- 645 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., Mamatjanov, 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., Ringer, 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.
- 650 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.
- 660

- 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., Erkkilä, 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.
- 665 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.
- 670 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.
- 675 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.
- 680 Kobayashi, S., Ota, Y., Harada, Y., Ebata, 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.
- 685 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.
- 690 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.
- 695 Malkki, P. and Tamsalu, R. E.: Physical Features of the Baltic Sea., 252, Finnish marine research, edn., 1985.

- 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.
- 700 Messenger, 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.
- Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Philipps, 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.
- 705 Mironov, D.: Parameterization of lakes in numerical weather prediction. Description of a lake model., Tech. Rep. 11, Deutscher Wetterdienst, 2008.
- 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., UCAR/NCAR, <https://doi.org/10.5065/D6RR1W7M>, 2013.
- 710 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., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Otlé, 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.
- 720 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.
- 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.
- 730

- 735 Seland, V., Bentsen, M., Olivie, 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.
- 740 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.
- 745 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.
- 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.
- 750 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.
- 755 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., Luysaert, 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.
- 760 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.
- 765 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.
- 770

- 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.
- 775 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-](https://doi.org/10.5194/hess-22-6279-2018)
- 780 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.

# Conclusion

## Sommaire

---

Synthèse . . . . .	135
Perspectives . . . . .	138
Amélioration de la paramétrisation . . . . .	138
Piste d'amélioration de FLake . . . . .	138
Implémentation des lacs dans le système de routage des rivières . . . . .	139
Prédictions futures . . . . .	139

---

## Synthèse

Les interactions entre le climat et les lacs sont nombreuses et conduisent à des effets notables. Les échanges d'énergie et d'eau entre l'atmosphère et les surfaces d'eau libres affectent directement la couche limite de l'atmosphère. On peut notamment évoquer le lissage des températures saisonnières de l'air ou encore l'augmentation des précipitations liée à la présence de lacs. Les lacs agissent aussi sur des cycles plus longs comme celui du carbone en relâchant du  $CO_2$  dans l'atmosphère et en le stockant dans les sédiments ou encore sur le cycle du méthane avec l'augmentation des émissions du fait de l'augmentation des températures globales. Le réchauffement climatique a aussi de nombreux impacts sur les lacs, comme l'augmentation de leur température, la variation de leur régime thermique, leur disparition ou encore la salinisation de certains lacs. Toutes ces interactions montrent l'intérêt de coupler les modèles de lacs avec les modèles de climat. Dans cette perspective, cette thèse visait à implémenter les lacs dans le modèle de surface continentale ORCHIDEE qui est une des composantes du modèle de climat de l'IPSL. Plus particulièrement, nous nous sommes focalisés sur la modélisation du bilan d'énergie qui est la première étape clé pour permettre de modéliser les autres interactions des lacs avec le climat. Nous nous sommes donc donné comme objectifs de choisir un modèle thermique de lac et de l'implémenter dans ORCHIDEE pour ensuite l'étalonner et le valider à l'aide de données satellitaires.

Une première introduction générale du bilan d'énergie terrestre a été effectuée avant d'aborder le cas plus spécifique des lacs pour tenir compte des différents aspects dans les développements du couplage. Plus particulièrement, nous avons décrit la structure thermique des lacs avec sa composition dont la couche de mélange caractérisée par une forte turbulence, la thermocline avec un gradient de température important et enfin l'hypolimnion qui caractérise les eaux les plus profondes avec une température constante. Les mécanismes liés au forçage météorologique



ont été évoqués pour comprendre leur impact sur la structure thermique du lac. On peut synthétiser cela par la compétition entre deux phénomènes avec le rayonnement solaire qui tend à stabiliser la colonne d'eau et le vent qui tend à augmenter la couche de mélange. Il existe une exception en hiver où le rayonnement solaire tend à créer des mouvements convectifs à cause d'une anomalie de la densité de l'eau où celle-ci est plus dense à  $4\text{ }^{\circ}\text{C}$  qu'à  $0\text{ }^{\circ}\text{C}$ . À partir de ces observations, plusieurs modèles ont été développés pour être utilisés dans les modèles de climat. Parmi ceux-là, le modèle FLake qui repose sur le concept d'auto-similarité du profil de température, a été choisi.

Les deux modèles, FLake et la partie énergétique d'ORCHIDEE, ont été décrits en détails avant d'expliquer la mise en œuvre de leur couplage. Nous avons choisi la possibilité d'agréger les lacs en plusieurs types et puis d'effectuer un bilan d'énergie pour chacun en maintenant le bilan d'énergie sur les surfaces végétalisées tel qu'il est réalisé dans ORCHIDEE.

Pour paramétrer correctement le modèle FLake couplé avec ORCHIDEE, nous avons mené une analyse de sensibilité sur les différents flux de chaleur turbulents et la température de surface simulés par FLake. Les paramètres pris en compte dans l'analyse sont la profondeur, les albédos et les coefficients d'extinction de l'eau, de la neige et de la glace, le fetch et un coefficient interne au modèle qui correspond au temps de relaxation de la thermocline. Nous avons étudié trois climats différents (tropical, tempéré et boréal) pour avoir un aperçu général. La méthode retenue pour mener cette analyse est celle de Sobol. Elle est basée sur la décomposition de la variance où chaque terme correspond à l'influence d'un paramètre ou au l'effet combiné de plusieurs. Cette méthode échantillonne l'espace des paramètres avec une approche de type quasi-Monte Carlo permettant d'obtenir une distribution uniforme des paramètres tirés au sort. Les intervalles d'incertitudes des paramètres ont été fixés à l'aide la littérature spécialisée. Pour la profondeur, nous avons choisi trois catégories où l'incertitude est calculée à l'aide des variations temporelles du niveau d'eau des lacs. Les lacs les moins profonds sont les plus sensibles avec une variance de température qui atteint  $2\text{ }^{\circ}\text{C}$  pour la température pour les périodes chaudes quel que soit le climat et de  $20\text{ W/m}^2$  et  $5\text{ W/m}^2$  pour le flux de chaleur latente et sensible. La profondeur et le coefficient d'extinction expliquent la majorité de la variance. Pour les lacs les plus profonds, la variance baisse en dessous de  $1\text{ }^{\circ}\text{C}$  pour la température et en dessous de  $10\text{ W/m}^2$  et  $2\text{ W/m}^2$  pour le flux de chaleur latente et sensible. L'albédo et le coefficient de relaxation de la thermocline expliquent en partie la variance observée sur la température de surface et les flux de surface. Cette analyse montre le comportement différent entre les lacs peu et très profonds, on recommande donc de les séparer dans les modèles de climats et c'est ce que nous avons choisi de faire dans la suite des développements.

La suite des travaux de la thèse a porté sur la validation du couplage d'ORCHIDEE et de FLake. Nous avons choisi la base de données GloboLakes pour valider les températures de surface des lacs et la base de données Global Lake and River

Ice Phenology Database pour valider la durée de gel simulée. Plusieurs configurations ont été testées pour analyser la sensibilité au forçage et la sensibilité à la paramétrisation de la profondeur. Les forçages utilisés ont une résolution de 0.5 degrés (WFDEI, CRUJRA, CRUNCEP) et 0.25 (ERA5, E2OFD) degrés. Nous avons comparé 5 stratégies de paramétrisation de la profondeur. Quatre configurations utilisent un seul type de lac et une dernière configuration représente les lacs en trois catégories en utilisant la médiane pour agréger la profondeur. Dans les quatre premières configurations, nous avons utilisé les moyennes arithmétique, harmonique et quadratique et la médiane pour agréger la profondeur des lacs au sein d'une maille. Nous avons finalement réalisé des simulations entre 2000 et 2016 avec 25 ans de spin-up. Nous avons observé une grande sensibilité au forçage. La médiane des distributions des erreurs quadratiques moyennes des simulations des forçages à 0.5 degrés se situe entre  $2.8^{\circ}C$  et  $3.2^{\circ}C$  et se situe autour de  $2.75^{\circ}C$  pour les forçages d'une résolution de 0.25 degrés. Nous avons néanmoins remarqué que la stratégie de paramétrisation de la profondeur est moins influente que le choix du forçage atmosphérique sur les températures de surface simulées. La différence notable porte sur la modélisation de la période de gel entre la configuration entre trois types de lac et un type de lac considéré. La configuration avec trois types de lac dégrade légèrement la modélisation de la période de gel. Dans une dernière partie du travail de la validation, nous avons comparé le modèle couplé avec les lacs et sans les lacs à l'échelle globale. Nous avons montré qu'ORCHIDEE simule des flux d'évaporation sur l'année plus importants quand il est couplé avec FLake sur la plupart des endroits du monde. En moyenne annuelle sur le monde en été, l'évaporation dans la simulation prenant en compte les lacs est plus forte de  $0.25\text{ mm/jour}$ , passant de  $0.82\text{ mm/jour}$  à  $1.07\text{ mm/jour}$ . Sur la période estivale, l'évaporation passe de  $1.5\text{ mm/jour}$  à  $1.72\text{ mm/jour}$ . Nos résultats montrent donc l'importance de bien prendre en compte ces surfaces dans les modèles globaux.

## Perspectives

### Amélioration de la paramétrisation

La paramétrisation du modèle peut être améliorée. Le second paramètre qui ressortait de l'analyse de sensibilité était le coefficient d'extinction. L'intérêt de spatialiser ce paramètre à l'échelle globale a été émis par [Heiskanen et al. \(2015\)](#). Ils ont notamment mis en évidence que les lacs turbides (au delà de  $0.5 m^{-1}$ ) sont peu sensibles à un changement du coefficient d'extinction alors que les lacs très clairs le sont. Des images satellites peuvent être utilisées pour créer une base de données de coefficient d'extinction. [Zolfaghari et al. \(2017\)](#) a appliqué l'algorithme CoastColors sur des images MERIS pour obtenir le coefficient d'extinction du lac Erie. Cet algorithme permet d'avoir d'autres propriétés optiques des lacs, il donne plus précisément le coefficient d'extinction pour 15 longueurs d'onde différentes. [Zolfaghari et al. \(2017\)](#) ont mis en évidence l'amélioration des simulations de FLake avec cet étalonnage, mais précisent que l'évolution du coefficient de manière mensuelle n'améliore pas les résultats des simulations de FLake à l'échelle globale. Avec cette étude, les auteurs montrent qu'il est possible d'exploiter les images satellitaires pour construire une base de données de ce coefficient à l'échelle globale.

Il serait intéressant d'étudier la sensibilité du modèle provenant de ce paramètre en fonction du changement climatique. En effet, le changement climatique peut agir sur la turbidité de l'eau. La structure thermique peut alors être changée et les brassages aussi. [Woolway et al. \(2017\)](#) a étudié les changements de régime induit par le réchauffement climatique, mais avec un coefficient d'extinction constant au cours des années. Tester des hypothèses comme l'augmentation du coefficient d'extinction avec l'augmentation des températures est pertinent pour pouvoir prédire les brassages, ce qui permet de faire des hypothèses sur l'état de l'oxygénation et donc de faune et de la flore du lac par exemple.

L'albédo est moins influent sur le bilan d'énergie de FLake, mais il est relativement facile d'améliorer sa paramétrisation dans le cadre actuel. Le modèle ORCHIDEE utilise des produits MODIS pour paramétrer l'albédo. Il est envisageable d'utiliser ces mêmes données pour préciser l'albédo des lacs et de le spatialiser de la même manière que la profondeur.

La paramétrisation de la profondeur peut être améliorée en prenant en compte les évolutions temporelles. La prochaine mission SWOT pourra permettre de contraindre les profondeurs à l'aide des observations satellitaires des niveaux d'eau des lacs.

### Pistes d'amélioration de FLake

Le modèle FLake lui-même peut aussi être amélioré. On peut noter trois pistes principales qui sont en discussion. La première est la modélisation de la structure saline pour les lacs concernés. En se basant sur le même concept qu'utilise FLake pour la structure thermique, on peut modéliser une couche thermohaline pour les lacs salins. Le principe reposerait sur le même concept que le profil de température où il y aurait une couche de mélange avec une concentration saline constante et puis

une thermohaline qui suivrait le même concept que la thermocline. Cependant, des inversions de densité ou de température peuvent apparaître dans ces profils comme la densité varie selon ces deux paramètres. Il faut donc rajouter d'autres équations pour lier ces profils.

La deuxième amélioration serait la modélisation de l'hypolimnion pour les lacs profonds. Une autre relation sur la hauteur de la thermocline doit être mise en équation. Une idée serait que cette hauteur soit proportionnelle à la hauteur de la couche de mélange divisée par le coefficient de Richardson en reprenant les paramétrisations de [Deardorff et al. \(1980\)](#). De plus, il faut trouver d'autres relations pour calculer la température au fond du lac et aussi à l'interface entre la thermocline et l'hypolimnion. Enfin, la dernière piste d'amélioration serait de paramétrer différemment la thermocline pour prendre une plus grande variété de profils thermiques. Cette idée est décrite par [Golosov et al. \(2018\)](#).

### **Implémentation des lacs dans le système de routage des rivières**

Le routage d'ORCHIDEE est décrit en détails par [Nguyen-Quang et al. \(2018\)](#). Le réseau est décrit à une résolution très fine (1 km) avec la base de données HydroSHEDS, mais il est agrégé à une résolution plus grande. Des graphes au sens mathématique représentent la connexion des mailles d'ORCHIDEE. Plusieurs bassins sont représentés au niveau de chaque noeud du graphe qui représente les écoulements de surface lents et rapides. Chaque bassin a un temps caractéristique différent pour céder une partie de l'eau qu'il contient au noeud aval du graphe. Une des perspectives est de coupler le lac dans le système de routage. Il faut d'abord modéliser chaque terme du bilan de masse du lac, c'est-à-dire l'évaporation, les précipitations, les échanges avec le souterrain, le ruissellement qui alimente le lac, les débits des rivières entrants et sortants. FLake calcule le débit d'évaporation. Il faut donc créer un bassin de type lac qui communique avec les autres réservoirs du routage. Des problèmes peuvent apparaître sur le manque de cohérence entre la manière d'agréger les lacs pour calculer le bilan d'énergie et celle de l'eau.

### **Prédictions futures**

Une des perspectives à plus long terme est de pouvoir mener des simulations couplées avec l'atmosphère, ce qui permettrait de faire des prédictions climatiques. Pour arriver à ce point, il faut d'abord faire le couplage entre LMDZ et ORCHIDEE-FLake.

Actuellement, du fait du couplage implicite de la résolution de la température, LMDZ gère qu'une seule colonne thermique d'ORCHIDEE dans le sol. Pour éviter des instabilités numériques, les profils de température de l'atmosphère et du sol sont calculés de manière implicite.

Il est prévu de développer une version d'ORCHIDEE où il est possible de faire un bilan d'énergie différent par type de surface (pas seulement que pour les lacs comme ici). Ensuite, il faut réaliser le couplage entre ces différents profils de température d'ORCHIDEE avec le profil de température de l'atmosphère calculé par LMDZ.

Il faudra donc vérifier la stabilité du modèle et sa convergence dans ce cas. Un autre problème persiste avec FLake. Des variables dans certains cas sont calculées avec un schéma numérique explicite comme la hauteur de la couche de mélange et les processus n'ont pas les mêmes temps caractéristiques. Cependant, ce temps est plus long pour les lacs, ce qui devrait faciliter le couplage. Un travail devra être fourni pour vérifier la stabilité et la convergence en introduisant FLake dans la version couplé de LMDZ avec ORCHIDEE.

Après ces développements techniques, l'analyse des rétroactions du climat sur les lacs pourra être étudiée. Le modèle du CRNM qui utilise SURFEX a estimé l'impact des lacs sur le modèle couplé avec l'atmosphère. [Le Moigne et al. \(2016\)](#) ont montré par exemple que sur les régions avec de nombreux lacs, l'humidité spécifique a augmenté. L'ajout des lacs dans un modèle couplé peut par exemple augmenter les précipitations et des orages comme le montre [Thiery et al. \(2015\)](#) sur le lac Victoria. Il serait donc intéressant de voir l'effet du réchauffement climatique avec l'ajout des lacs sur la couche limite de l'atmosphère. On peut par exemple faire l'hypothèse que les régions humides avec un nombre de lacs élevé pourront faire face un nombre encore plus élevé d'orages avec le réchauffement climatique. Un tel couplage avec l'atmosphère permettrait aussi de quantifier le ralentissement du réchauffement climatique des régions avec de nombreux lacs puisque ceux-ci absorberaient la chaleur comme le montre l'étude de [Vanderkelen et al. \(2020\)](#) sur le siècle dernier. Un dernier exemple serait de quantifier le changement de régime thermique de lacs comme [Woolway and Merchant \(2019\)](#) l'ont fait sur un nombre restreint de lacs.

## Annexe

### Article 3

Dans cet article, nous nous sommes focalisés sur l'apport potentiel des observations futures de la mission SWOT sur les modèles de lacs qui simule le bilan d'énergie plus spécifiquement sur le modèle Flake. En effet SWOT nous permet de connaître l'évolution des surfaces d'eau. D'après l'analyse de sensibilité menée dans cette thèse, nous savons que la profondeur est un paramètre important pour simuler le bilan d'énergie avec Flake. Nous avons vérifié l'apport de SWOT sur les différents lacs réservoirs de la Seine et sur une gravière dont l'élévation du niveau d'eau est suivi in situ.

Dans la première partie, j'ai participé à la génération des erreurs simulées sur les niveaux d'eau sur les différents lacs et sur la région des gravières pour les différentes dates de passages du satellite. Dans la deuxième partie, j'ai participé à la réalisation des différentes simulations Flake. Nous avons considéré trois cas :

- Il n'y a pas de données donc on prend la profondeur maximale du barrage de chaque lac réservoir.
- On actualise les profondeurs à chaque passage du satellite
- On actualise la profondeur à chaque passage avec l'erreur associée qui a été préalablement simulée. Les profondeurs ont été préalablement fournies au pas de temps journalier par les gestionnaires des grands lacs réservoirs de la Seine.

Nous avons pu notamment constater la différence sur les températures de surface et les flux entre les différentes simulations. Ainsi, SWOT a un réel potentiel pour améliorer les résultats des simulations Flake en contraignant la profondeur des lacs.

Article

# Characterization of SWOT Water Level Errors on Seine Reservoirs and La Bassée Gravel Pits: Impacts on Water Surface Energy Budget Modeling

Catherine Ottlé <sup>1,\*</sup> , Anthony Bernus <sup>1</sup>, Thomas Verbeke <sup>1</sup>, Karine Pétrus <sup>1</sup>, Zun Yin <sup>1</sup>, Sylvain Biancamaria <sup>2</sup>, Anne Jost <sup>3</sup> , Damien Desroches <sup>4</sup>, Claire Pottier <sup>4</sup>, Charles Perrin <sup>5</sup>, Alban de Lavenne <sup>5</sup>, Nicolas Flipo <sup>6</sup>  and Agnès Rivière <sup>6</sup> 

<sup>1</sup> Laboratoire des Sciences du Climat et de l'Environnement (LSCE), CNRS-CEA-UVSQ, F-91190 Gif-sur-Yvette, France; anthony.bernus@lsce.ipsl.fr (A.B.); thomas.verbeke@lsce.ipsl.fr (T.V.); karine.petrus@lsce.ipsl.fr (K.P.); zun.yin@lsce.ipsl.fr (Z.Y.)

<sup>2</sup> Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS), UPS (OMP-PCA), F-31400 Toulouse, France; sylvain.biancamaria@legos.obs-mip.fr

<sup>3</sup> CNRS, EPHE, METIS, Sorbonne Université, F-75005 Paris, France; anne.jost@sorbonne-universite.fr

<sup>4</sup> Centre National d'Études Spatiales, F-31400 Toulouse, France; damien.desroches@cnes.fr (D.D.); claire.pottier@cnes.fr (C.P.)

<sup>5</sup> INRAE, UR HYCAR, Université Paris Saclay, F-92160 Antony, France; charles.perrin@inrae.fr (C.P.); alban.de-lavenne@inrae.fr (A.d.L.)

<sup>6</sup> Centre de Géosciences, MINES Paris-Tech, F-77305 Fontainebleau, France; nicolas.flipo@mines-paristech.fr (N.F.); agnes.riviere@mines-paristech.fr (A.R.)

\* Correspondence: catherine.ottle@lsce.ipsl.fr; Tel.: +33-169-086-268

Received: 27 July 2020; Accepted: 4 September 2020; Published: 8 September 2020



**Abstract:** The Surface Water and Ocean Topography (SWOT) space mission will map surface area and water level changes in lakes at the global scale. Such new data are of great interest to better understand and model lake dynamics as well as to improve water management. In this study, we used the large-scale SWOT simulator developed at the French Space National Center (CNES) to estimate the expected measurement errors of the water level of different water bodies in France. These water bodies include five large reservoirs of the Seine River and numerous small gravel pits located in the Seine alluvial plain of La Bassée upstream of the city of Paris. The results show that the SWOT mission will allow to observe water levels with a precision of a few tens of centimeters (10 cm for the largest water reservoir (Orient), 23 km<sup>2</sup>), even for the small gravel pits of size of a few hectares (standard deviation error lower than 0.25 m for water bodies larger than 6 ha). The benefit of the temporal sampling for water level monitoring is also highlighted on time series of pseudo-observations based on real measurements perturbed with the simulated noise errors. Then, the added value of these future data for the simulation of lake energy budgets is shown using the FLake lake model through sensitivity experiments. Results show that the SWOT data will help to model the surface temperature of the studied water bodies with a precision better than 0.5 K and the evaporation with an accuracy better than 0.2 mm/day. These large improvements compared to the errors obtained when a constant water level is prescribed (1.2 K and 0.6 mm/day) demonstrate the potential of SWOT for monitoring the lake energy budgets at global scale in addition to the other foreseen applications in operational reservoir management.

**Keywords:** SWOT errors; Seine reservoirs; La Bassée gravel pits; FLake model; radar altimetry

## 1. Introduction

### 1.1. On the Key Role of Lakes and Reservoirs on the Water and Energy Cycles

The continuous growth of human pressure on freshwater resources especially under global warming requires a comprehensive monitoring of the hydrological cycle. Since surface water is the main source of human water uses [1,2], an explicit quantification of the water storage variations in reservoirs and lakes is the prerequisite to improve the management of surface water resources. Moreover, even if the inland water bodies cover only 3.5% of the Earth's non glaciated terrestrial area [3], their impact on the atmosphere can be significant. For instance, in the high latitudes of the northern hemisphere, water surfaces can account for a large fraction of the land surface and can play a significant role on the regional weather and climate [4]. Indeed, the specificities of the radiative and thermal properties of water and its capacity to freeze under cold weather conditions impact the energy and water transfers with the atmosphere and the resulting local meteorology and climatology. The impacts of artificial reservoirs on the terrestrial hydrological cycle also appeared to be quite important, not only by affecting the river regimes but also by substantially contributing to freshwater withdrawals through the increased evaporation and the promotion of irrigation linked to the increased water availability [5,6]. As an example, it was demonstrated how the large increase of reservoirs in China during the last 20 years have impacted regional water budgets, especially evapotranspiration and its long-term trends [7]. In the same part of the world, a simple dam operation model was developed to explain the time variability of the discharges of the Yellow river and to separate the effects of climate change from human infrastructures [8].

Because of the key role of lakes and reservoirs in the continental hydrological cycle and associated feedbacks, it is crucial to represent them in land surface models in order to better understand and model their various contributions to the different components of the water budget. For that purpose, the modeling of both energy and water balances (strongly coupled through the evaporation) is a requirement as far as the interactions with the atmosphere need to be explicitly described. Such modeling implies the representation of both surface extent and storage of the lakes/reservoirs considered as well as observations for the parameter calibration. Such measurements can be provided by in situ gauges generally available for managed reservoirs but these data are often insufficient, especially considering global modeling. Therefore, alternative approaches based on space remote sensing appear very promising to fill this gap.

### 1.2. On the Promising Use of Satellite Imagery for Lakes and Reservoir Monitoring at the Global Scale

Satellite optical imagery such as the multispectral radiometers onboard the Sentinel-2, SPOT, or Landsat platforms, can map water bodies at global scale with a temporal sampling of a few days and at high spatial resolution up to a few meters. These devices have been used by various studies to provide global maps of wetlands and water bodies, such as the most recent products of [3,4,9]. Moreover, satellite nadir altimetry may complement these data by providing water levels along the satellite track for some rivers and lakes, as demonstrated by [10–12] for floodplains and summarized by [13]. The combination of both types of measurements allowed to estimate water storage changes in lakes and reservoirs, as shown by various studies, e.g., [14].

In a very near future, wide swath altimetry will allow to move a step forward by providing both water elevation above a reference surface and water area thanks to Synthetic Aperture Radar (SAR) interferometry technics. Such instruments will be able to continuously observe water storage changes with a high spatial and temporal resolution and looks very relevant to monitor surface water storage changes and discharges.

The new Surface Water and Ocean Topography (SWOT) satellite mission developed jointly by the U.S. National Aeronautics and Space Administration (NASA) and the French Centre National d'Etudes Spatiales (CNES), with contributions from the Canadian and the UK Space Agencies (CSA and UKSA), with a launch planned for 2021, will provide such data. This satellite will carry a Ka-band



radar interferometer able to measure water surface elevation over oceans and continental surfaces with an intrinsic horizontal resolution of a few meters for water bodies larger than 0.06 km<sup>2</sup> with a repeat cycle of 21 days [15]. It will have a vertical accuracy of at least 10 cm after aggregating pixels (with no vegetation) over 1 km<sup>2</sup> of water [16]. Over lakes and reservoirs, the SWOT mission is expected to provide simultaneously a time varying water mask and the corresponding surface water elevation. Several studies have been devoted to the preparation of these new data, and two instrument simulators were developed to simulate the expected SWOT measurements. First of all, the SWOTsim simulator [17] was developed by NASA/Jet Propulsion Laboratory (JPL) to simulate the expected SWOT measurements based on realistic satellite error and orbit characteristics. For example, SWOTsim was used to evaluate temporal and spatial errors in global water reservoirs [18], and to explore SWOT future capacity to monitor storage change and outflow in 20 reservoirs of the Mekong river basin [19]. It was used also on smaller water bodies by [20] to assess the potential of SWOT to monitor water volumes in Sahelian ponds and lakes. On the French side, CNES developed a simplified large-scale simulator [21] to simulate the expected SWOT errors, given the foreseen sources of errors and their statistical properties. Our study is one of the first to use this tool.

### 1.3. Study Objectives

The aim of this study is to assess the contribution of the future SWOT mission to the monitoring of the water levels of the Seine reservoirs and gravel pits of the alluvial plain upstream the Paris city region for the assessment of evaporation, or more largely, surface energy balance. For that purpose, we first propose to use the CNES large scale SWOT simulator to characterize the sampling frequency and to assess the future instrument errors expected on these water bodies. The simulated errors are used in a second step in a lake model to quantify the impacts of these uncertainties on the energy budget and especially on the water evaporation and surface temperature. Our objective is, therefore, to answer the following questions:

- What will be the spatial and temporal sampling of SWOT on the Seine reservoirs and alluvial gravel pits and the expected errors on water level and volume measurements?
- How these errors translate in model uncertainties on surface temperature and fluxes using a state of the art and widely used lake model?

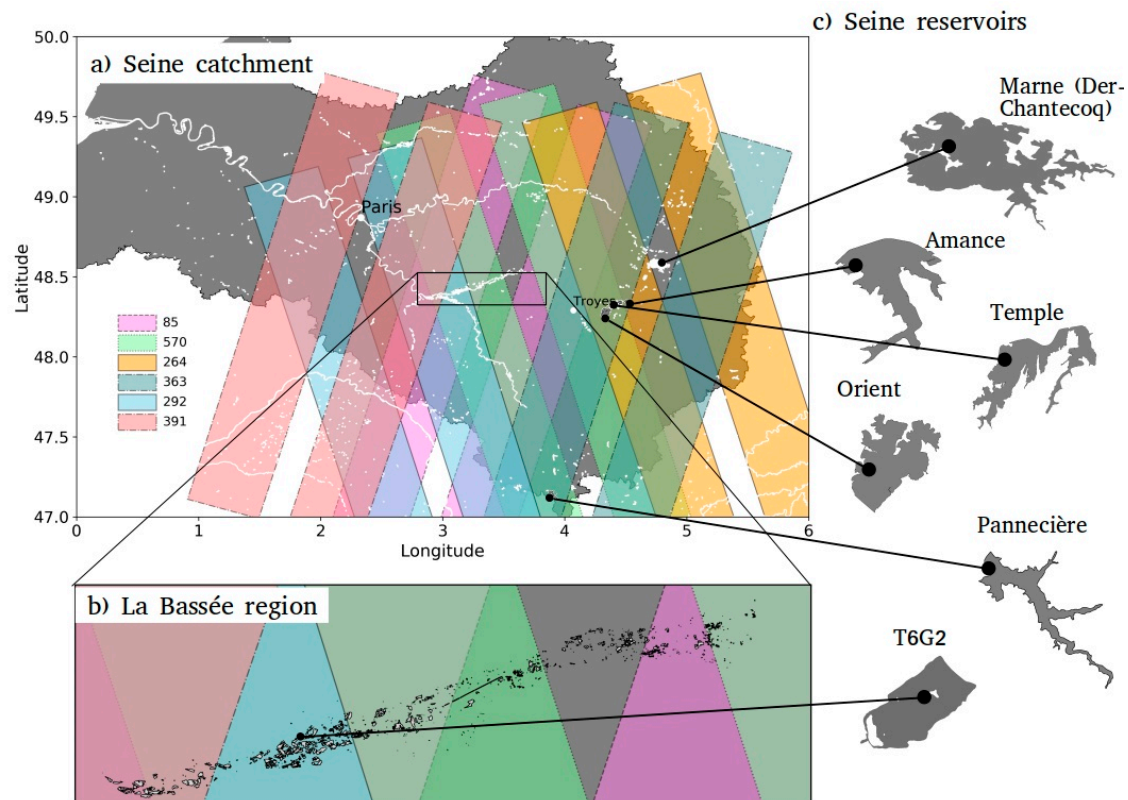
## 2. Study Region and SWOT Coverage

### 2.1. Seine Reservoirs and La Bassée Gravel Pits

The study region, the location of the Seine reservoirs and La Bassée gravel pits considered in the following are presented in Figure 1. Table 1 lists the reservoirs name, surface area, water holding capacity and maximal depth.

We focus here on the upstream part of the Seine river, which is one of the major rivers of France (777 km long), flowing through the city of Paris into the English Channel and draining a basin of about 79,000 km<sup>2</sup>. The river is regulated for discharge control (reducing flood peaks and holding water levels during low-flow periods), with five reservoirs built on the main stream and tributaries (the Marne, Yonne and Aube rivers) managed by the Seine Grands Lacs basin authority. Among these hydraulic structures, the oldest one is the Pannecièrre dam, built on the Yonne tributary and commissioned in 1949, which covers an area of 5.2 km<sup>2</sup>, with a maximum water holding capacity of 80 Mm<sup>3</sup> and a maximum depth of 49 m. The four other structures are all diversion reservoirs connected to the Seine, the Marne and the Aube rivers, which operate in the following way: the water is removed from the river by diversion in winter and spring, to fill the reservoir and reduce the floods downstream; storage is released to the river during summer and early fall to maintain water levels and hold the production of drinking water and water needs for crops irrigation [22,23]. Their surface area varies from 5 km<sup>2</sup> (Amance on the Aube river, which is the smallest, with a water holding capacity of 22 Mm<sup>3</sup>) to 48 km<sup>2</sup> (Der-Chantecoq on the Marne river, which is the largest, with a holding capacity of 349 Mm<sup>3</sup>).

In between the confluences of the Aube-Seine and Yonne-Seine rivers, the La Bassée alluvial plain is a regularly flooded section of the Seine river of about 364 km<sup>2</sup> (about 70 km long and 5 km wide) connected to a chalk aquifer characterized by the presence of numerous gravel pits (about thousand identified, among which more than 530 present a surface area less than 1 ha). Some of them are monitored, such as the gravel pit “T6G2” equipped with thermal and water pressure instruments (A. Jost, personal communication). These instruments are designed to monitor the water level and temperature, in order to better understand the water transfers between the chalk and alluvial aquifers and the gravel pit.



**Figure 1.** The Surface Water and Ocean Topography (SWOT) swaths over the Seine reservoirs and over the La Bassée region. The rivers and water bodies were extracted from the cartographic Carthage database (<https://geo.data.gouv.fr/fr/datasets>). The location of the five reservoirs studied and T6G2 gravel pit is highlighted. The same color code is used for the two maps.

**Table 1.** Key parameters of the five reservoirs of the Seine river and the La Bassée zone studied.

Seine Lakes	Der-Chantecoq	Orient	Temple	Amance	Pannecièrre	La Bassée
Area (km <sup>2</sup> )	48	23.2	18	5	5.2	28
Holding capacity (Mm <sup>3</sup> )	349	208	148	22	80	-
Maximal depth (m)	20	25	22	22	49	12
River controlled	Marne	Seine	Aube	Aube	Yonne	Seine
Number of observations over the 21-day cycle	0	3	3	2	2	4
SWOT observation days	-	9, 13, 20	9, 13, 20	9, 13	13, 20	3, 10, 14, 20
Orbit numbers	-	264, 363, 570	264, 363, 570	264, 363	363, 570	85, 292, 391, 570

## 2.2. SWOT Mission and Orbits

SWOT will be equipped with a Ka-band (35.75 GHz or 8.6 mm wavelength) SAR interferometer instrument (called KaRIn) with near nadir incidence angles (between 0.6° and 3.9°), which will provide measurements of water surface elevation over ocean and continental surfaces with a pixel resolution

of 6 m in the along-track direction and 10 m to 60 m from near range to far range in the across-track direction, depending on the location within the two 50-km wide swaths provided by the two antennas on each side of the satellite with a 20 km gap in between [15]. On continental surfaces, SWOT will allow to measure water level changes for wetlands, lakes and reservoirs larger than  $250 \times 250 \text{ m}^2$  with a repeat cycle of 21 days. However, because of the swath coverage and the orbit characteristics, the number of observations per cycle will vary according to the latitude. For our studied reservoirs and gravel pits, the number of revisits can be up to four observations within 21 days, with time gaps ranging from 4 to 17 days within this revisit cycle.

The days of observation per revisit period are listed for all the reservoirs and the La Bassée region, in Table 1. As an example, Figure 1 shows the sampling area of all the orbits which will intersect the Seine reservoirs and the La Bassée alluvial plain. Unfortunately, one reservoir (Marne-Der-Chantecoq) within the Marne watershed, located at the intersection of the nadir gap areas, will never be observed by the SWOT-KaRIn instrument. Nevertheless, six SWOT orbits will sample the studied area (85, 264, 292, 363, 391 and 570) and their respective swath is visualized in Figure 1 (note that the color code associated to each orbit will be the same hereafter). The Orient and Temple reservoirs will be sampled three times over the 21-day cycle repeat orbit pass, at 9-, 13- and 20-day regular intervals by the orbits 264, 363 and 570. The Amance and Pannecièrè reservoirs will be sampled twice at 9-, 13- (respectively, 13-, 20-) day intervals at orbits 264 and 363 (respectively, 363, 570). The La Bassée region will be sampled up to four times at 3-, 10-, 14- and 20-day intervals at orbits 85, 292, 391 and 570.

### 3. Lake Model and SWOT Simulator

#### 3.1. FLake Lake Model

The FLake lake model [24] was chosen to assess the potential contribution of the near future SWOT data to the monitoring of the energy budgets of the water bodies of the Seine basin for environmental studies. FLake is a one-dimensional (1D) model dedicated to the simulation of the surface-atmosphere energy transfers and principally designed to be coupled to weather prediction or climate modeling. The model considers two layers in the lake: a top mixed layer with uniform vertical temperature and a thermocline below, parameterized according to a self-similarity concept. A sediment layer can be optionally added at the bottom, based on the same concept. Freezing processes and evolution of a snowpack layer above the ice layer are also simulated with a bulk approach, and a parametric representation of the temperature profile. The model is already implemented in many numerical weather prediction and climate models [25–27] to cite but a few, and was chosen to represent water surfaces in the ORCHIDEE land surface model, the continental part of the Institut Pierre Simon Laplace (IPSL) climate model [28].

Two energy budgets (for the mixed layer and for the bulk lake) are calculated to solve the temperature profile and calculate the energy surface fluxes (radiative and turbulent fluxes) at the lake-atmosphere interface, given the atmospheric forcing (incoming solar and atmospheric radiations, air temperature and humidity and wind speed). In order to characterize the thermal and radiative properties of the lake, four main parameters need to be prescribed: the lake depth, the water surface albedo, the light extinction coefficient and the wind fetch. A preliminary study [28] quantified the respective sensitivity of the model outputs to these four parameters. It confirmed the key role of lake depth on the simulation of the surface temperature and fluxes, and therefore, the expected contribution of SWOT instrument for improving the model predictions.

#### 3.2. SWOT Simulator

The SWOT large scale simulator [21] developed by CNES was used here to generate the water surface elevation (WSE) errors on the studied water bodies. The simulator is composed of two parts: one module to simulate the satellite orbitography and the instrument swaths given the characteristics of the mission (therefore, the revisit times and the geometric configuration), and a second module

to calculate the final error budget of the SWOT measure at the pixel level. The simulator accounts for the two main sources of error that will affect the KaRIn measurements, i.e., the random errors corresponding to the instrument thermal noise and geometric signal decorrelation and the systematic errors linked mostly to tropospheric propagation delay and rolling errors. The random errors are generated at pixel level and can be approximated as white noises with normal distribution and standard deviations varying with the pixel location in the swath. Unlike the instrumental noise, the systematic errors are constant in the swath for a given time. They are modeled for each date as random variables in this version of the simulator. Here, the systematic errors are generated with Gaussian white noise with a standard deviation equal to 0.05 m. Gaussian white noise is also used to generate the instrumental errors, with a standard deviation varying around 3 m for viewing angle less than 3° and increasing up to 30 m for larger values. Thanks to the random nature of the signal and the large number of measurements within the studied lakes, the instrument budget error may be largely reduced by averaging over a number of pixels and then dividing by the square root of this number [29,30]. It should be noted that the topographic and vegetation layover are not accounted for in this simplified simulator, but a geolocation error due to the interferometric baseline roll angle can be estimated. It has been activated here in the water surface elevation (WSE) error simulations. Additionally, dark water pixels corresponding to the case of specular water reflectivity were not accounted for in the simulations. More information on the main features of the large-scale simulator can be found in [21].

#### 4. Simulation of SWOT-Like WSE Errors

##### 4.1. Method

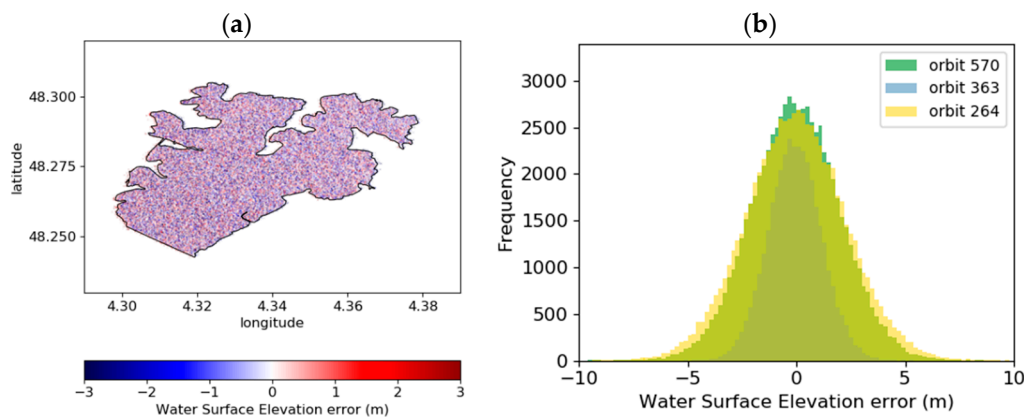
WSE errors were simulated with the large-scale SWOT simulator on the studied Seine reservoirs and La Bassée gravel pits. The contours of these water bodies as georeferenced polygons were extracted from the CARTHAGE database (French national hydraulic database BD CARTHAGE® IGN; see [www.sandre.eaufrance.fr/Referentiel-hydrographique](http://www.sandre.eaufrance.fr/Referentiel-hydrographique)) and redrawn from the 2014 ortho-photography database (BD ORTHO® IGN). The simulations were performed for a five-year period to mimic the expected lifetime of the satellite set as a goal [16]. We analyzed the WSE errors at the pixel scale and at the lake scale after averaging the random errors on our studied water bodies. The spatial and temporal variations are analyzed according to the satellite orbit for the four Seine reservoirs and 1062 gravel pits of La Bassée region.

##### 4.2. Results

###### 4.2.1. WSE Errors on the Seine Reservoirs

###### Pixel-Scale WSE Error

Figure 2 presents an example of the pixel cloud obtained for one pass (orbit 264, day 9 of the simulation) over the Orient reservoir. The noisy nature of the signal is clearly seen with pixel error values ranging between  $-10.09$  m and  $+10.03$  m, leading to a mean error value equal to  $-0.024$  m. The effect of the geolocation error is clearly seen on the western part of the reservoir with larger positive errors. This can be explained by the fact that the geolocation error impacts differently positive and negative vertical errors, in such a way that positive errors will be moved away from the satellite track and the other way round for negative errors. The distribution of the errors calculated over the five-year simulation (Figure 2b) shows the lowest range of variation for the orbit 363 compared to the orbits 570 and 264 (presenting the largest errors). This is an illustration of the impact of the distance of the Orient reservoir to the nadir line, which is roughly 20 km, 45 km and 47.5 km for the orbits 363, 570 and 264, respectively (see Figure 1). Therefore, the viewing angle is indeed lower for the orbit 363, explaining that the errors range only between  $-5.34$  m and  $+4.85$  m, in this case.



**Figure 2.** SWOT-like errors simulated for the Orient reservoir: (a) pixel cloud of the errors simulated for one date (Orbit 264, day 9 of the simulation); (b) histogram of the pixel errors for the three orbits covering the Orient reservoir calculated over one repeat cycle (i.e., for days 9, 13 and 20 of the simulation).

#### Lake-Scale WSE Error

When averaged over the lake surface area and because of the random nature of the signal, the pixel errors compensate and the mean WSE error (as it will be observed by the SWOT-KaRin instrument at the lake scale) decreases drastically down to a few centimeters. The standard deviation (std) and maximal (absolute value) WSE errors calculated over the five-year simulation period (90 cycles) are summarized in Table 2 for the four reservoirs and the instrumented gravel pit of La Bassée (T6G2). The surface extent and the average number of SWOT-like measurements ( $N_p$ ) for each orbit configuration are also provided. Note that  $N_p$  can vary slightly from one pass to the other, due to the geolocation error. It can be seen that for all the lakes, the standard errors are very small, and up to 0.11 m for the smallest water body (T6G2). The maximum errors are of a few tens of centimeters and up to 0.32 m for the Pannecière reservoir and T6G2 gravel pit. The location of the water body inside the swath has a limited impact. The lower the incidence angle, the lower the error, as shown for example by the difference of about 20% between the errors obtained for the Orient reservoir for the Orbit 264 and those obtained for the Amance reservoir. In summary, these simulations clearly show that the studied water bodies are large enough to allow the measurement of their depth with a decimetric precision, even for a small lake of about 20 ha as the gravel pit T6G2.

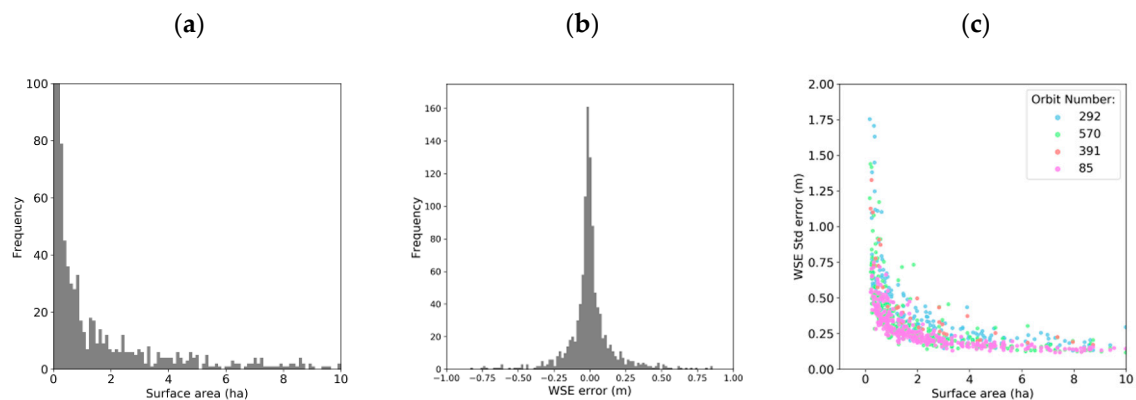
**Table 2.** Standard (Std) and maximal (Max) SWOT-like water surface elevation (WSE) errors (in meters) simulated for the four studied reservoirs of the Seine and the T6G2 gravel pit. Surface extent in km<sup>2</sup> and average number of SWOT pixels ( $N_p$ ) are given for each water body.

Seine Lakes	Orient			Temple			Pannecière			Amance			T6G2		
Area (km <sup>2</sup> )	23			18			5.2			5			0.22		
$N_p$	Max (m)	Std (m)		$N_p$	Max (m)	Std (m)	$N_p$	Max (m)	Std (m)	$N_p$	Max (m)	Std (m)	$N_p$	Max (m)	Std (m)
Orbit 264	72,643	0.3	0.11	30,755	0.27	0.09	-	-	-	10,366	0.28	0.09	-	-	-
Orbit 570	68,403	0.28	0.10	40,660	0.22	0.09	7405	0.32	0.09	-	-	-	528	0.28	0.14
Orbit 363	34,611	0.28	0.10	13,707	0.3	0.10	6083	0.26	0.10	3450	0.22	0.08	-	-	-
<b>Orbit 85</b>	-	-	-	-	-	-	-	-	-	-	-	-	321	0.32	0.12
Orbit 292	-	-	-	-	-	-	-	-	-	-	-	-	680	0.26	0.11

#### 4.2.2. WSE Errors on La Bassée Gravel Pits

In the region of La Bassée, 1062 gravel pits were extracted from the CARTHAGE database. Their size ranges between 35 m<sup>2</sup> and 0.9 km<sup>2</sup> (90 ha) and the size distribution presented in Figure 3a clearly shows the dominance of the small gravel pits with size lower than 1 ha. The simulator has been run on all the gravel pits sampled by the four orbits covering the region, i.e., on 801 water bodies,

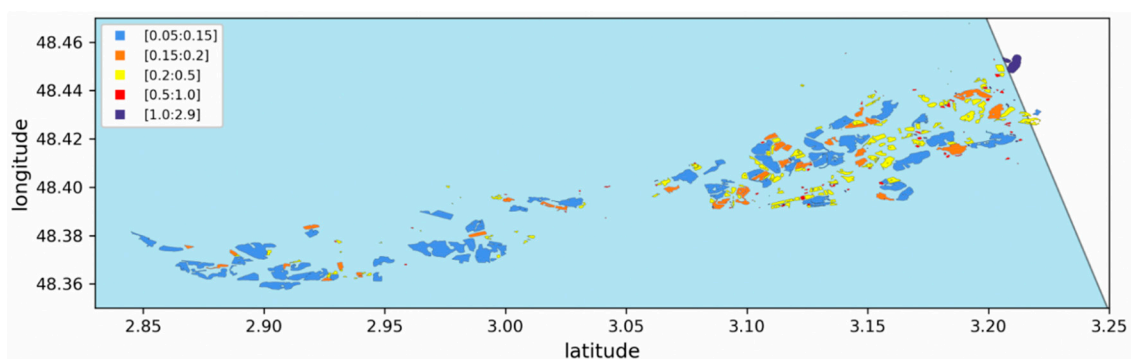
and on a five-year period. Given the 21-day orbit cycle, the number of observations over five years is therefore equal to 90.



**Figure 3.** (a) Gravel pits size distribution; (b) Water Surface Elevation (WSE) errors simulated for 801 gravel pits of La Bassée on a five-year period; (c) standard deviation of the simulated SWOT-like WSE errors according to the surface area and for the four orbits sampling the 801 gravel pits studied.

The simulated SWOT-like errors averaged at the gravel pit scale are displayed on Figure 3b. The errors are larger than the ones simulated for the Seine reservoirs and range mostly in the  $[-1$  m,  $+1$  m] interval. Among all the errors simulated on the whole period (136,214 errors in total), around 18% are above 0.5 m (30% being lower than 0.1 m).

The impact of the size of the gravel pit on the standard deviation of the measurement error is clearly shown on the scatterplot presented in Figure 3c, where the smaller gravel pits show the largest standard deviation errors, up to 1.75 m. The four orbits show the same features with an exponential decrease to a value lower than 0.25 m when the surface area exceeds 4 ha. Among all, the orbit 85 presents the lowest values and the orbit 292 the largest. This can be explained by the location of the La Bassée region within the different swaths, with the orbit 85 observing the region with the lowest viewing angles (lowest distance to the nadir track). In any case, the results show that for the gravel pits that will be observed operationally by SWOT (with a size larger than 6 ha), the precision of WSE will be better than 0.25 m.



**Figure 4.** Standard deviation of the SWOT-like errors simulated on the gravel pits of the western part of the La Bassée region and for the Orbit 292 for the five-year period studied.

The spatial analysis of the WSE errors on the whole set of simulations did not show any correlation with the size of the gravel pit ( $\text{Corr} = -0.08$ ), but some features appeared when the analysis was reduced to one orbit. As an illustration, Figure 4 presents the standard deviation of the errors calculated for all the gravel pits sampled by the orbit 292 located on the western part of the La Bassée region (including the T6G2 gravel pit), and calculated over the 90 SWOT-like observations of the five-year study period. Indeed, the plot shows larger values for the smaller lakes and larger values when moving

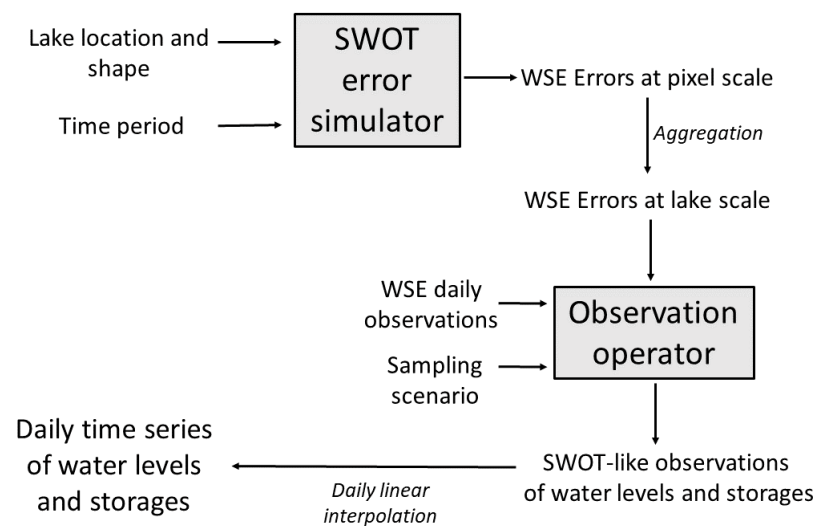
from west to east, i.e., moving to the far edge of the swath. The largest gravel pits show standard values lower than 0.15 m, which is the order of magnitude obtained for T6G2 (Table 2). Moreover, most of the gravel pits (70%) present standard values lower than 0.5 m and only 11% show values larger than 1 m. The largest value (2.83 m) is obtained for a gravel pit that is located on the border of the swath and is strongly affected by geolocation errors.

#### 4.2.3. Contribution of SWOT in WSE Daily Monitoring

Given the expected SWOT WSE errors, we found it interesting to assess the impacts of these errors on the daily monitoring of our studied water bodies according to the period of revisit, in order to quantify the benefit of SWOT temporal sampling on the reconstruction of daily records of both WSE and water volumes.

- Methodology

For this purpose, we have generated, for each of the Seine reservoirs, SWOT-like observations of surface water level by adding the simulated errors to in-situ observations provided by the Seine Grands Lacs basin authority over the period 2003–2007 at daily time step. It should be noted that these data are water levels provided by in-situ probes or observers. These pseudo-observations were linearly interpolated at a daily time step, to generate daily time series of surface water levels as they could be derived from SWOT observations. The corresponding water volumes were estimated using empirical relationship between water storage and water level, calibrated on each lake with actual observations as proposed by [31]. In order to assess the impact of temporal sampling on the WSE errors, we generated different daily time series corresponding to different sampling scenarios, considering either a single measurement over the 21-day cycle period with each orbit separately or all the possible measurements (two or three depending on the reservoir studied). The latter scenario will be called “All orbits” hereafter. The developed methodology is displayed in Figure 5.

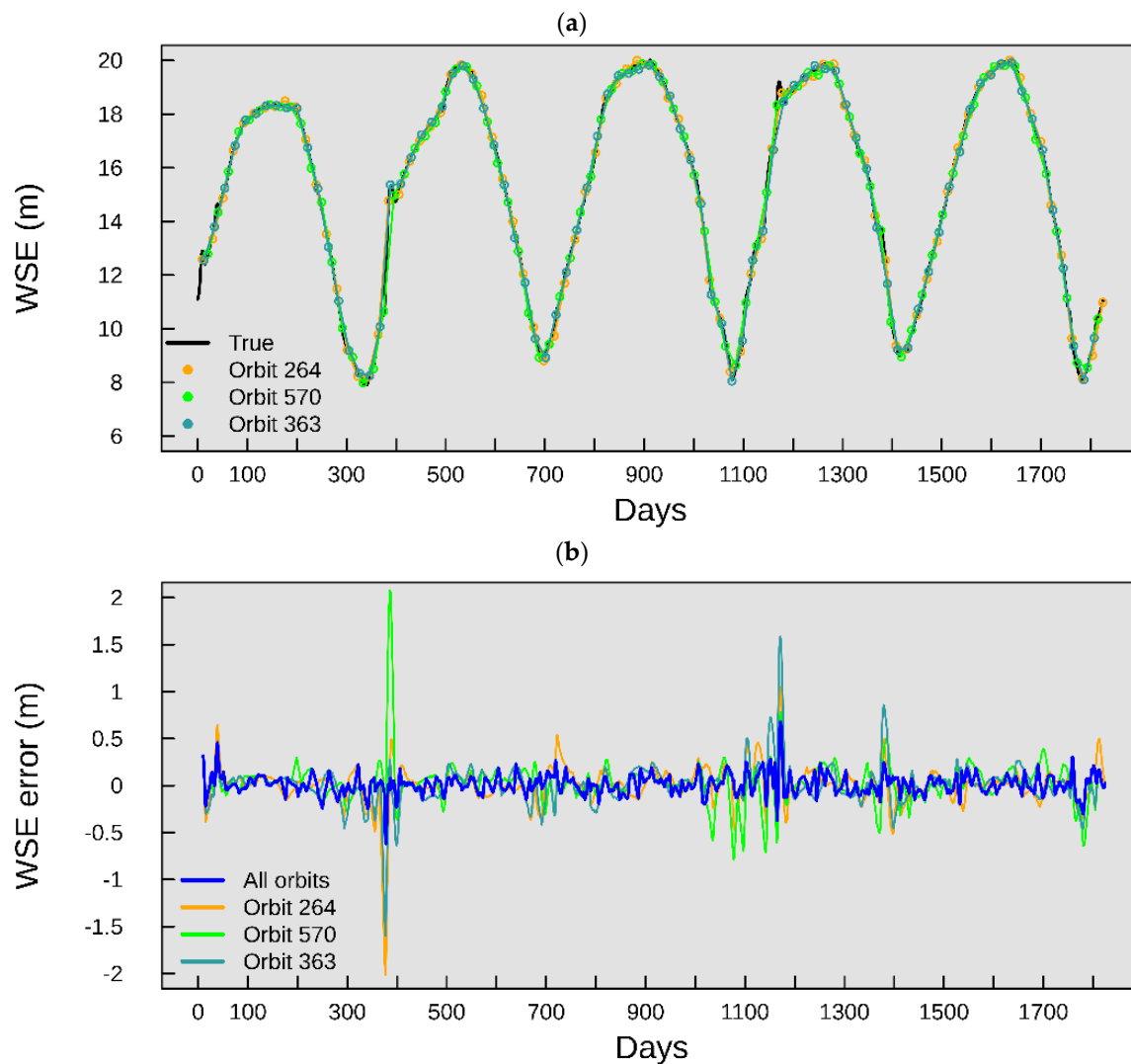


**Figure 5.** Methodology workflow used to generate the daily time series of water levels and storages for the reservoirs and gravel pits studied.

- Results

As an example, Figure 6a shows the daily variation of WSE as it could be derived from SWOT observations on the Orient lake on a five-year period considering the three orbits covering the lake, therefore, for four sampling scenarios (each of the three orbits individually and all jointly). The true observations are plotted on the same graph as a reference. The WSE errors calculated as the differences between the daily time series and the true observations for the four sampling scenarios, are plotted on

Figure 6b. The benefit of the more frequent sampling is clearly seen with a significant reduction of the errors when all the orbits are considered jointly (std = 0.1 m). In comparison, the errors are two to four times larger when a single orbit is considered. As expected, the errors are larger in winter, when the temporal variations of the reservoir volumes are the largest, because of the management operations (water released to river to empty the reservoirs before the floods and then stored from November on). During these periods, the WSE anomalies can reach several meters, like in winter 2003–2004 (Year 2 of the simulation, Figure 6b), when it goes up to 2 m in the case where only orbit 570 is used to follow water level.



**Figure 6.** Water surface elevation (a) and anomalies (b) simulated for the Orient lake, considering three orbits separately and together, in the period 2003–2007.

Table 3 summarizes the values of the maximal errors and standard deviations of the WSE errors, calculated for the four reservoirs on a five-year period. The largest error is found for the Orient lake and orbit 570 (WSE error of 2.08 m, corresponding to a storage error of 4.6 Mm<sup>3</sup>). In all cases, the errors are lower when all the observations are used to monitor the water level and storage, but the reduction of error is larger for the Orient and Temple reservoirs compared to the Pannecièrre and Amance ones. Indeed, the error is reduced by a factor of three for the first two reservoirs and by only 40% and 20% for Pannecièrre and Amance, respectively. This result can be easily explained by the more frequent sampling of Orient and Temple compared to the other two reservoirs (three observations/cycle instead of two). It shows that the improvement increases with the number of revisits, and also that two

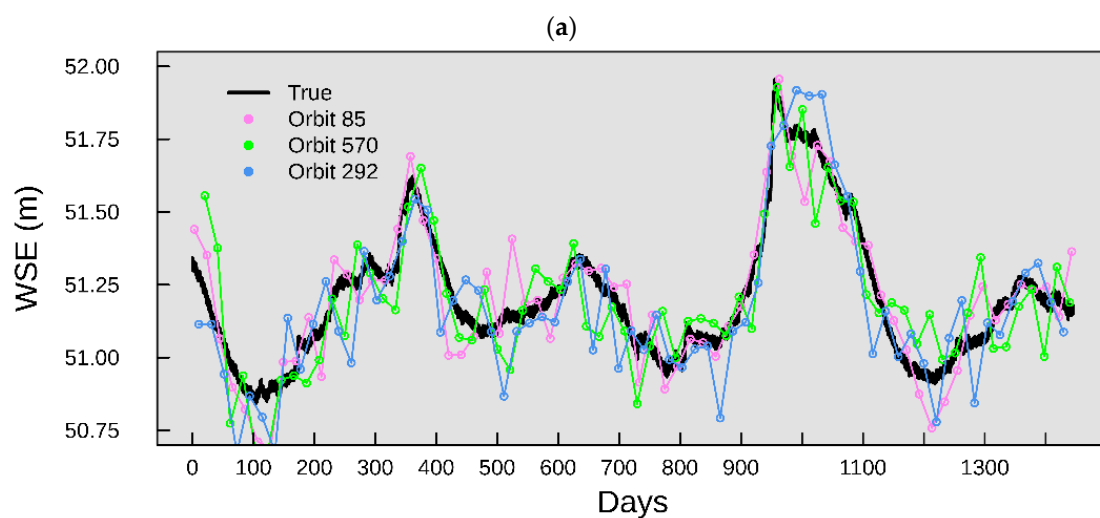


observations well distributed in the 21-day interval (Pannecière case) are better than when they are clustered (Amance case), especially during filling/emptying time periods when the WSE varies rapidly.

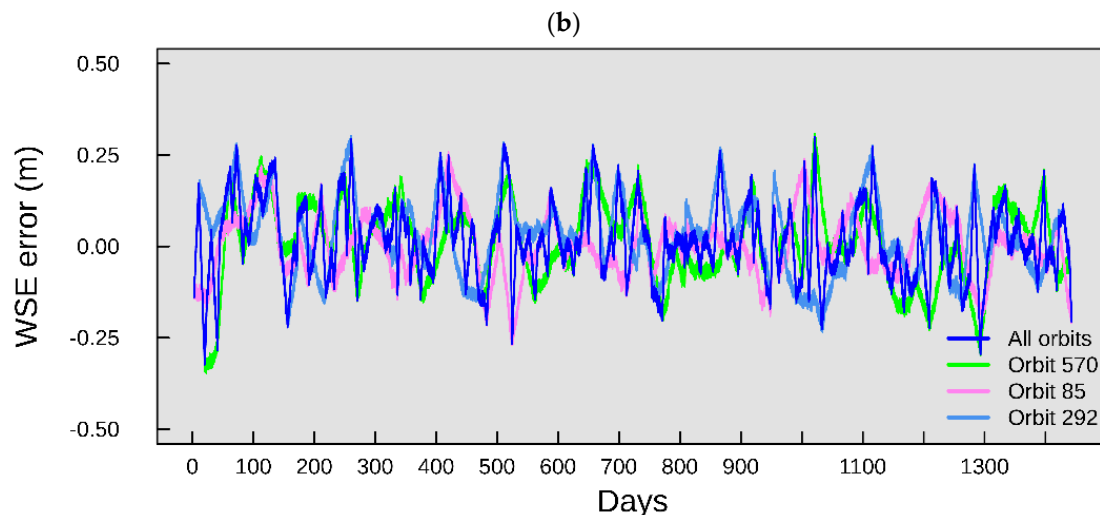
**Table 3.** Maximal error and standard deviations of the errors on the estimated daily WSE (in meters) for each individual orbit and the maximum sampling scenario (all the orbits). The corresponding water storage errors (in  $\text{Mm}^3$ ) were derived for the four gauged reservoirs. Results for the T6G2 gravel pit are also listed.

Seine Lakes	Orient		Temple		Pannecière		Amance		T6G2	
	Max(m) ( $\text{Mm}^3$ )	Std(m) ( $\text{Mm}^3$ )	Max(m) ( $\text{Mm}^3$ )	Std(m) ( $\text{Mm}^3$ )	Max(m) ( $\text{Mm}^3$ )	Std(m) ( $\text{Mm}^3$ )	Max(m) ( $\text{Mm}^3$ )	Std(m) ( $\text{Mm}^3$ )	Max(m)	Std(m)
Sampling scenario	2.02	0.21	1.46	0.22	-	-	1.58	0.23	-	-
Orbit 264	4.4	0.46	1.08	0.16	-	-	0.47	0.07	-	-
Orbit 363	1.6	0.21	1.49	0.25	1.99	0.37	1.65	0.23	-	-
	3.5	0.46	1.1	0.18	2.3	0.44	0.49	0.07	-	-
Orbit 570	2.08	0.22	1.43	0.25	2.04	0.38	-	-	0.35	0.11
	4.6	0.48	1.06	0.18	2.45	0.45	-	-	-	-
Orbit 85	-	-	-	-	-	-	-	-	0.27	0.09
Orbit 292	-	-	-	-	-	-	-	-	0.3	0.1
All	0.68	0.1	0.7	0.11	1.49	0.22	1.38	0.17	0.32	0.1
	1.5	0.22	0.52	0.08	1.79	0.26	0.41	0.05	-	-

The same analysis has been performed for the monitored gravel pit T6G2, for which four years of measurements of water elevation above sea level are available on the period 13 May 2015–12 June 2019. As an indication, the elevation of the surrounding area is around 53 m. In the case of an unmanaged lake, the WSE temporal variations (Figure 7a) are much smoother than those for the managed reservoirs previously described. Therefore, the daily maximal and standard errors are much smaller (up to a factor of seven for the maximal error and a factor of two for the standard error) and the benefit of increased sampling is less significant (Figure 7b and Table 3). The anomalies show about the same patterns whatever the sampling scenario. The maximum error is equal to 0.35 m with a standard deviation of 0.11 m (in the case of orbit 570) with no reduction when all orbits are used. This result would suggest that for such small water bodies, the SWOT error is too large and cannot be reduced by an increase of the sampling up to three observations per cycle; a larger sampling would be needed in that case.



**Figure 7.** Cont.



**Figure 7.** Daily evolution of the water surface elevation (a) and anomalies (b), both in meters, calculated for the T6G2 gravel pit on a four-year period (2015–2019), considering each orbit separately and together.

## 5. Impacts on Water Energy Budget Modeling

The contribution of SWOT data on the estimation of the energy budget of the Seine reservoirs and gravel pits has been assessed through sensitivity experiments performed with the FLake lake model [24]. As demonstrated by [28] through a global sensitivity analysis, lake depth plays a key role in the calculation of lake temperature and heat fluxes, and consequently, in the assessment of lake evaporation. In order to quantify the added value of SWOT measurements on the monitoring of these surface variables, we have designed a set of modeling experiments on our studied water bodies, corresponding to various levels of knowledge of the key depth parameter. The study was performed on the four Seine reservoirs as well as on the T6G2 gravel pit.

### 5.1. FLake Simulation Experiment Design

Three depth scenarios were built to quantify the contribution of SWOT data on surface temperature and heat fluxes modeling:

- The first scenario (called REF) considers that the actual depth of the reservoir/gravel pit is known precisely and provided by daily local measurements; FLake simulations forced by such true data will serve as reference.
- In the second scenario (called MAX), the water level is not measured but the maximal capacity of the reservoir/gravel pit is known (as given by the literature) and will be used to prescribe a constant value to the depth parameter.
- In the third scenario (called SWOT), the water depth is regularly measured from space (and prone to errors) and updated at the SWOT sampling rate considering all the orbits covering the lake. It is assumed here that the water depths will have be derived from the water surface elevation measured by SWOT provided a reference measurement.

For these three scenarios, we generated depth time series for the five Seine water bodies (four reservoirs and T6G2 gravel pit) on the two time periods previously studied (2003–2007 for the reservoirs and 2015–2019 for T6G2). All the simulations were performed using the same atmospheric conditions (incident radiation, air temperature and humidity, wind speed) based on atmospheric reanalysis provided by the WFDEI (WATCH Forcing Data methodology applied to ECMWF Re-Analysis-Interim) database [32]. In order to isolate the contribution of a better knowledge of the lake depth, all the other parameters prescribed in FLake were kept the same, equal to the standard values proposed by [24] and available at <http://www.flake.igb-berlin.de>. It should be noted that the sediment module of FLake

was not used in these simulations and that for all the simulations, a three-year spin-up period was performed to reach model equilibrium and get rid of incorrect initialization states. The results are discussed in terms of surface temperature ( $T_S$ ) and latent ( $H_L$ ) and sensible ( $H_S$ ) heat fluxes, which are the main lake variables interacting with the atmosphere above.

## 5.2. Modeling Results

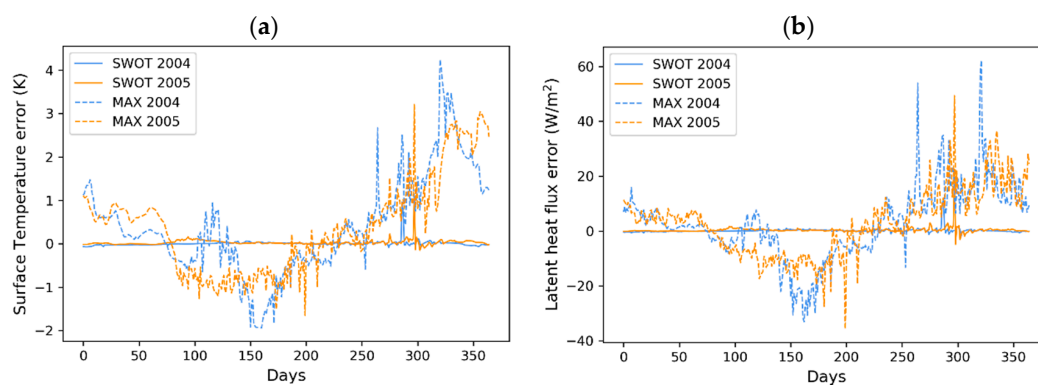
The FLake simulations based on the MAX and SWOT scenarios were compared to the one forced with true depth values (scenario REF). The differences can be interpreted as errors propagated from the depth errors to the surface temperature and fluxes of the studied lakes.

### 5.2.1. Model Errors on Seine Reservoirs

The analysis was performed on the four studied reservoirs and is summarized in Table 4, in terms of maximal and standard deviation errors. Figure 8 shows the daily surface temperature and evaporation flux errors obtained on the Orient reservoir on two years, 2004 and 2005, for the SWOT and MAX scenarios.

**Table 4.** Standard (Std) and maximum (max) errors on FLake surface temperature (in K) and surface heat fluxes (in  $W/m^2$ ), calculated over a five-year period, for the four reservoirs of the Seine and T6G2 gravel pit. The comparison is done for the MAX and SWOT simulations in comparison to REF.

Seine lakes	Orient		Temple		Amance		Pannecière		T6G2	
Max Depth (m)	25		22		22		49		6	
Scenario 2 (MAX)										
	Max	Std	Max	Std	Max	Std	Max	Std	Max	Std
$T_S$ error (K)	5	1.2	6.4	1.1	6.9	1.2	6.7	1.2	5	0.9
$H_S$ error ( $W/m^2$ )	56.3	9.2	60.2	7.8	66.2	8.8	64.6	8.8	49.3	5.7
$H_L$ error ( $W/m^2$ )	88.4	14.4	96.3	12.1	104.4	14.7	103.3	15.6	81.9	10.8
Scenario 3 (SWOT)										
	Max	Std	Max	Std	Max	Std	Max	Std	Max	Std
$T_S$ error (K)	3.3	0.2	5.2	0.4	1.92	0.11	3.4	0.5	6.6	0.4
$H_S$ error ( $W/m^2$ )	37.7	1.4	59.3	3.1	16.6	0.81	28.1	3.6	77.4	3.9
$H_L$ error ( $W/m^2$ )	61	2.7	91.4	4.6	19.3	0.97	56.2	6.6	152.6	5.1



**Figure 8.** Temporal evolution of the surface temperature (in K) (a) and latent heat flux (in  $W/m^2$ ) (b) errors calculated for the Orient reservoir and for the years 2004 and 2005. The two depth scenarios SWOT (depth forced to SWOT-like measurements) and MAX (constant depth equal to reservoir maximal capacity) are compared to the REF (actual depth) one.

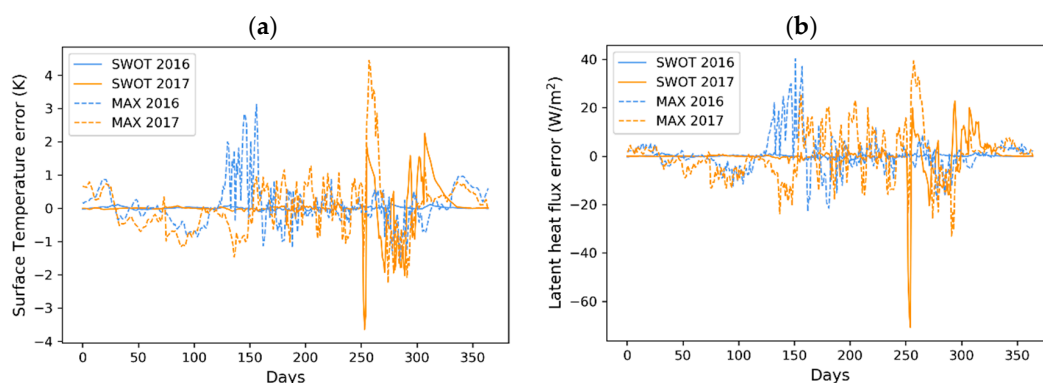
The plots clearly show the strong reduction of the errors in the SWOT simulations, on  $T_S$  and also  $H_L$ . The systematic overestimation of  $T_S$  in winter season and underestimation in summer is the result of the overestimation of the reservoir depth by the MAX scenario (larger in winter, as discussed previously) compared to the true values, as shown in Figure 6a. Such depth errors, which can be

up to 17 m at the end of fall in the case of the Orient lake, can drastically modify the lake regime, the convection flows and the temperature profile. Indeed, the largest errors occur during the periods when the reservoir level is minimum at the end of fall and in summer when the atmospheric demand is the largest, depth errors having much lower impacts in winter, when the potential evaporation is low. Here, during these two years,  $T_S$  errors can exceed 4 K, which translate in  $H_L$  errors up to  $60 \text{ W/m}^2$ , namely an evaporation error of  $2.5 \text{ mm/day}$ . In the SWOT scenario, the depth errors are substantially reduced with maximum values up to  $0.68 \text{ m}$  (see Table 3), which will lead to  $T_S$  and  $H_L$  errors less than 1 K and  $10 \text{ W/m}^2$ , respectively.

Table 4 summarizes the results obtained for the four reservoirs, and shows the benefit of the SWOT future observations on the reservoir energy budget monitoring. Maximal and standard deviation errors are reduced in all cases. The standard errors as compared to the MAX scenario are reduced by a factor of 6 for the Orient lake and up to 10 in the case of the Amance lake, which is the one presenting the largest error in the MAX scenario. The fluxes errors (latent heat  $H_L$  and sensible heat  $H_S$  fluxes) are reduced by the same factors as expected.

### 5.2.2. Model Errors on T6G2 Gravel Pit

On the much shallower T6G2 gravel pit, whose depth spatially varies up to 6 m, the SWOT contribution to the reduction of energy budget errors is also significant. Figure 9 presents the  $T_S$  and  $H_L$  errors obtained for the two scenarios MAX and SWOT, simulated for the years 2016 and 2017 (for which the depth records cover the whole year). In this case, the depth translated errors are larger than for the deeper reservoirs, for at least two reasons: the depth errors are larger and the sensitivity of the energy budgets components to the depth parameter is larger for shallow lakes, as shown by [28]. The errors are dominant in spring and fall, which correspond to the two annual periods of lake mixing and the disappearance of the thermocline layer. The date of this turnover period is strongly dependent on the lake depth: a positive/negative error on the depth will lead to a time difference (delay/ahead, respectively) in the mixing process. Interestingly, these periods are well simulated both in 2016 and in spring 2017 in the SWOT scenario, whereas, in 2017, the fall turnover after DOY 250 seems to be delayed by a few days in both simulations MAX and SWOT. This is explained when looking to the SWOT errors simulated on the same period, which happened to overestimate the WSE of the gravel pit (see Figure 7a, around day 980), and consequently, lead to a delay of the lake turnover and an overestimation of  $T_S$  (since the temperature of the surface layer is larger than the bottom temperature at the end of summer). The larger magnitude of the SWOT errors on this small water body does not allow to monitor precisely this specific event. The errors statistics on the four-year period summarized in Table 4 highlight again the added value of the SWOT observations even for this small lake, with a reduction of the standard deviation of the errors by a factor of around two as compared to the MAX scenario, both on surface temperature and fluxes.



**Figure 9.** Temporal evolution of the surface temperature (a) and latent heat flux (b) errors, respectively, in K and  $\text{W/m}^2$ , calculated for the T6G2 gravel pit and for the two years 2016 and 2017. The two depth scenarios SWOT and MAX are compared to REF.

## 6. Discussion

### 6.1. SWOT Reduction of Errors

The SWOT mission will provide, in a very near future, surface elevation of water bodies with an unrivalled decimetric precision and with a repeat cycle of 21 days as a minimum. These new observations will help monitoring surface water fluxes and storages in rivers, lakes, reservoirs and wetlands. In this study, we focused on the Seine basin and its large reservoirs and alluvial plain upstream the city of Paris, and tried to assess the future contribution of these new observations to follow water levels but also surface fluxes through a lake energy budget model.

The large-scale simulator [21] was used to characterize the expected instrument errors on the studied water bodies. Based on the satellite orbit and swath characteristics, the sample rates and days of revisit were computed over four out of the five Seine reservoirs, (one of the reservoirs will unfortunately never be observed because of the instrument observation nadir gap) and 801 gravel pits of the alluvial plain. Some of the gravel pits will be sampled up to four times within the 21-day orbit cycle, which will provide measurements at time intervals lower than 6 days. The SWOT simulator results show that the expected measurements over the large reservoirs will be obtained with a precision around 0.10 m and up to 0.25 m for the smaller gravel pits with size down to 6 ha. These results are in line with the instrument budget errors [29] and demonstrate the tremendous interest of SWOT future data for water management purposes, especially for poorly gauged and ungauged reservoirs and lakes. They are also in agreement with the previous works of [18–20] based on the JPL-SWOTsim simulator, even though our errors appear slightly larger. Thus, the Agoufou lake in Sahel, with an area varying from 100 ha to 250 ha, [20] showed that WSE could be retrieved with an accuracy better than 4 cm.

These errors were used in a second step to quantify water levels and storages errors in an objective of daily monitoring, using all the SWOT observations potentially available. Daily time series were generated by linear interpolation of SWOT observations and were compared to true measurements, in order to assess the potential of SWOT for the daily monitoring of the Seine reservoirs. The results show a slight degradation of the WSE estimation with large errors up to 2 m, occurring when only one measurement in the 21-day interval is available and during the periods of filling/emptying of the reservoirs when the water level can change very rapidly. Fortunately, these errors are considerably reduced if the reservoir is sampled more frequently. In the case of the Orient reservoir, the errors are reduced by a factor of up to 3 when the three observations within the 21-day period are used. For smaller and unmanaged water bodies such as the Seine gravel pits, the error degradation linked to the linear interpolation is less significant and the benefit of frequent revisit compared to only one was not demonstrated, because of the smoother variations of WSE and the larger errors, which would benefit from a more frequent revisit.

Finally, the time series of WSE were used in the FLake model to assess the contribution of SWOT for lake energy budget modeling. In such models, the lake water level or mean depth is a very sensitive parameter, because it determines key processes such as the flow convection within the lake due to the water density gradients, which further affects the seasonal thermal stratification. In order to quantify the foreseen added value of SWOT, we generated SWOT-like time series of daily lake depth by direct assimilation of SWOT-like observations and used them to drive the FLake model. The results were compared to FLake simulations performed either with true daily observations of depth or using a constant value. The analysis highlighted the benefit of the SWOT observations for the modeling of surface water temperature and fluxes with, again, major error reductions. Therefore, this confirms the importance for developing tools and models to prepare the assimilation of the future SWOT products in catchment monitoring systems.

### 6.2. SWOT Perspectives for Reservoir Monitoring

This study was limited to a small number of water bodies, but their diversity, from the small gravel pit to the large reservoir, provided insights on the future SWOT errors on such water bodies (managed

and unmanaged). Furthermore, we addressed how these errors translate into surface temperature and flux errors in a lake energy budget modeling. The next step will be to use the SWOT data to also control the mass budget, through the assimilation of both the surface extent and level changes, or the use of the future SWOT products of lake water storage change, as shown by [33]. We anticipate that the use of such data will bring lake modeling major steps forward. Moreover, if we consider the added value of the future discharge products, it can be expected that the SWOT mission will bring new and very valuable data for hydrology.

The characterization of the SWOT errors was investigated using the large-scale simulator, which integrates the various types of errors expected after launch of the mission. In the simulation, the lake contour (and therefore its size) was kept constant. This assumption should not have impacted our error estimations on the Seine reservoirs because of their large size. It should be the same for most of the gravel pits whose size changes little during the year. Nevertheless, it will be interesting in the following, to run the JPL simulator on our studied water bodies to quantify the simulator uncertainties and also to test the last version of the large-scale simulator, which include some new features, such as spatialized tropospheric errors and improved rolling errors.

SWOT future lake products will include both water extent and water levels. The seasonal dynamics of water extent will also be a valuable information to better constrain lake and hydrological models. The global capacity of the instrument will allow to provide these data at global scale. Such information will be very useful to better assess the role of water bodies and wetlands on the global water cycle and more generally the land–atmosphere feedback for environmental and climate studies.

## 7. Conclusions

We used the CNES large scale SWOT simulator to simulate the expected instrument errors over the Seine reservoirs and gravel pits of its alluvial plain, in order to assess the potential of these new data to be used in the daily monitoring of water levels and storages. The pseudo-observations generated were used, firstly, to explore how SWOT will observe the studied reservoirs and the expected uncertainties, and secondly, to quantify the resulting surface temperature and fluxes errors in lake energy budget modelling.

Applied to four large reservoirs (with area varying from 5 km<sup>2</sup> to 23 km<sup>2</sup>) and 801 gravel pits of various sizes (areas varying from 35 m<sup>2</sup> to 0.9 km<sup>2</sup>), the simulator was used to estimate the expected precision of the future SWOT data (around 0.1 m for the reservoirs and better than 0.25 m for the smaller gravel pits). It was also used to provide the first quantitative assessment of the contribution of these data for constraining the simulation of the surface water energy budgets.

These findings confirm how valuable space altimetry is for reservoir/lakes monitoring, especially for ungauged water bodies or to complement in situ networks. The global capacity of SWOT to observe both water levels and surface extent and its increased spatial resolution and measurement precision, compared to the present instruments, will definitely help to better understand and model lakes and reservoirs dynamics. These new observations will certainly improve our knowledge about the role of water bodies on regional water and energy budgets, their impacts on the main biogeochemical cycles and their fate in global change scenarios.

**Author Contributions:** Conceptualization, C.O.; Methodology, C.O. and S.B.; Software, S.B., D.D., C.P. (Claire Pottier), and A.B.; Formal Analysis, C.O.; Resources, C.O., C.P. (Charles Perrin) and A.J.; Writing—Original Draft Preparation, C.O.; Writing—Review and Editing, C.O., A.B., S.B., A.J., D.D., C.P. (Charles Perrin) and A.d.L.; Visualization, T.V., A.B., K.P. and Z.Y.; Supervision, C.O.; Project Administration, C.O.; Funding Acquisition, C.O., C.P. (Charles Perrin), N.F., A.R. and A.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the French Space Agency (Centre National d'Etudes Spatiales) through the TOSCA-SPAWET program in preparation for the SWOT space mission. CEA (Commissariat de l'Énergie Atomique et aux Énergies Alternatives) is also acknowledged for funding part of A.B.'s PhD grant.

**Acknowledgments:** The Seine Grands Lacs basin authority is thanked for providing detailed data on the reservoirs of the Seine basin, as well as the French PIREN Seine program, who contributed to fund the data acquisition on the La Bassée gravel pits data. The anonymous reviewers are also acknowledged for their useful comments and suggestions on the previous version of this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Hanasaki, N.; Yoshikawa, S.; Pokhrel, Y.; Kanae, S. A global hydrological simulation to specify the sources of water used by humans. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 789. [CrossRef]
- Wada, Y.; Wisser, D.; Bierkens, M.F.P. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth Syst. Dyn.* **2014**, *5*, 15–40. [CrossRef]
- Verpoorter, C.; Kutser, T.; Seekell, D.A.; Tranvik, L.J. A global inventory of lakes based on high-resolution satellite imagery. *Geophys. Res. Lett.* **2014**, *41*, 6396–6402. [CrossRef]
- Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [CrossRef]
- Vorosmarty, C.J. The storage and aging of continental runoff in large reservoir systems of the world. *Ambio* **1997**, *26*, 210–219.
- Haddeland, I.; Heinke, J.; Biemans, H.; Eisner, S.; Flörke, M.; Hanasaki, N.; Stacke, T. Global water resources affected by human interventions and climate change. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3251–3256. [CrossRef]
- Mao, Y.; Wang, K.; Liu, X.; Liu, C. Water storage in reservoirs built from 1997 to 2014 significantly altered the calculated evapotranspiration trends over China. *J. Geophys. Res. Atmos.* **2016**, *121*, 10097–10112. [CrossRef]
- Yin, Z.; Otlál, C.; Ciaia, P.; Zhou, F.; Wang, X.; Jan, P.; Dumas, P.; Peng, S.; Piao, S.; Li, L.; et al. Irrigation, damming and streamflow fluctuations of the Yellow River. *Hydr. Earth Syst. Sci. Discuss.* **2020**, 1–28. [CrossRef]
- Sheng, Y.; Song, C.; Wang, J.; Lyons, E.A.; Knox, B.R.; Cox, J.S.; Gao, F. Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sens. Environ.* **2016**, *185*, 129–141. [CrossRef]
- Calmant, S.; Seyler, F. Continental surface waters from satellite altimetry. *Comptes Rendus Geosci.* **2006**, *338*, 1113–1122. [CrossRef]
- Kouraev, A.V.; Zakharova, E.A.; Samain, O.; Mognard, N.M.; Cazenave, A. Ob’river discharge from TOPEX/Poseidon satellite altimetry (1992–2002). *Remote Sens. Environ.* **2004**, *93*, 238–245. [CrossRef]
- Frappart, F.; Seyler, F.; Martinez, J.M.; León, J.G.; Cazenave, A. Floodplain water storage in the Negro River basin estimated from microwave remote sensing of inundation area and water levels. *Remote Sens. Environ.* **2005**, *99*, 387–399. [CrossRef]
- Crétau, J.-F.; Abarca-del-Río, R.; Bergé-Nguyen, M.; Arsen, A.; Drolon, V.; Clos, G.; Maisongrande, P. Lake volume monitoring from space. *Surv. Geophys.* **2016**, *37*, 269–305. [CrossRef]
- Crétau, J.F.; Jelinski, W.; Calmant, S.; Kouraev, A.; Vuglinski, V.; Bergé-Nguyen, M.; Maisongrande, P. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Adv. Space Res.* **2011**, *47*, 1497–1507. [CrossRef]
- Biancamaria, S.; Lettenmaier, D.P.; Pavelsky, T.M. The SWOT mission and its capabilities for land hydrology. In *Remote Sensing and Water Resources*; Springer: Cham, Switzerland, 2016; pp. 117–147.
- Desai, S. Surface Water and Ocean Topography Mission (SWOT) Project, Science Requirements Document. NASA/JPL technical document D-61923; 2018. Available online: [https://swot.jpl.nasa.gov/docs/D-61923\\_SRD\\_Rev\\_B\\_20181113.pdf](https://swot.jpl.nasa.gov/docs/D-61923_SRD_Rev_B_20181113.pdf) (accessed on 31 March 2020).
- Peral, E.; Rodríguez, E.; Moller, D.; McAdams, M.; Johnson, M.; Andreadis, K.; Arumugan, D.; Williams, B. *SWOT Simulator Quick User Guide*; NASA: Washington, DC, USA, 2016; p. D-79123.
- Solander, K.; John, C.; Reager, T.; Famiglietti, J.S. How well will the Surface Water and Ocean Topography (SWOT) mission observe global reservoirs? *Water Resour. Res.* **2016**, *52*, 2123–2140. [CrossRef]
- Bonnema, M.; Hossain, F. Assessing the Potential of the Surface Water and Ocean Topography Mission for Reservoir Monitoring in the Mekong River Basin. *Water Resour. Res.* **2019**, *55*, 444–461. [CrossRef]
- Grippa, M.; Rouzies, C.; Biancamaria, S.; Blumstein, D.; Crétau, J.; Gal, L.; Robert, E.; Gosset, M.; Kergoat, L. Potential of SWOT for Monitoring Water Volumes in Sahelian Ponds and Lakes. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2541–2549. [CrossRef]
- Desroches, D.; Pottier, C.; Blumstein, D.; Biancamaria, S.; Poughon, V.; Fjortoft, R. Large Scale Pixel Cloud Simulator and Hydrology Toolbox. In Proceedings of the SWOT Science Team Meeting, Montreal, QC, Canada, 23 June 2018.
- Villion, G. Rôle des lacs-réservoirs amont: Les grands lacs de Seine (The role of upstream dams: The large reservoirs of the Seine basin). *La Houille Blanche* **1997**, *8*, 51–56. [CrossRef]

23. Rizzoli, J.L.; Gache, F.; Durand, P.Y.; Jost, C. Flood management in Ile de France, toward a shared and global strategy. *La Houille Blanche-Revue Internationale de l'Eau* **2011**, *2*, 5–13. [[CrossRef](#)]
24. Mironov, D.V. *Parameterization of Lakes in Numerical Weather Prediction: Description of a Lake Model*; COSMO Technical Report No. 11; DWD: Offenbach, Germany, 2008.
25. Mironov, D.; Heise, E.; Kourzeneva, E.; Ritter, B.; Schneider, N.; Terzhevik, A. Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environ. Res.* **2010**, *15*, 218–230.
26. Balsamo, G.; Salgado, R.; Dutra, E.; Boussetta, S.; Stockdale, T.; Potes, M. On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus A Dyn. Meteorol. Oceanogr.* **2012**, *64*, 15829. [[CrossRef](#)]
27. Le Moigne, P.; Colin, J.; Decharme, B. Impact of lake surface temperatures simulated by the FLake scheme in the CNRM-CM5 climate model. *Tellus A Dyn. Meteorol. Oceanogr.* **2016**, *68*, 31274. [[CrossRef](#)]
28. Bernus, A.; Otlé, C.; Raoult, N. Variance based sensitivity analysis of FLake lake model for global land surface modeling. *J. Geophys. Res. Atmos.* **2020**. in revision.
29. Biancamaria, S.; Andreadis, K.M.; Durand, M.; Clark, E.A.; Rodriguez, E.; Mognard, N.M.; Oudin, Y. Preliminary characterization of SWOT hydrology error budget and global capabilities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *3*, 6–19. [[CrossRef](#)]
30. Esteban Fernandez, D. SWOT Project, Mission Performance and Error Budget. NASA/JPL technical document, D-79084. Available online: [https://swot.jpl.nasa.gov/docs/SWOT\\_D79084\\_v10Y\\_FINAL\\_REVA\\_06082017.pdf](https://swot.jpl.nasa.gov/docs/SWOT_D79084_v10Y_FINAL_REVA_06082017.pdf) (accessed on 1 April 2020).
31. Terrier, M. Flow Naturalization Methods and Uncertainty Estimates. Ph.D. Thesis, INRAE, Antony/AgroParisTech, Paris, France, 2020.
32. Weedon, G.P.; Balsamo, G.; Bellouin, N.; Gomes, S.; Best, M.J.; Viterbo, P. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.* **2014**, *50*, 7505–7514. [[CrossRef](#)]
33. Munier, S.; Polebistki, A.; Brown, C.; Belaud, G.; Lettenmaier, D.P. SWOT data assimilation for operational reservoir management on the upper Niger River Basin. *Water Resour. Res.* **2015**, *51*. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





## Bibliographie

- Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G. A., and Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, 54(6part2) :2283–2297.
- Alsdorf, D. E., Rodríguez, E., and Lettenmaier, D. P. (2007). Measuring surface water from space. *Reviews of Geophysics*, 45(2) :RG2002.
- Andrew J., T., Norman D., Y., Keller, B., Girard, R., Heneberry, J., Gunn, J. M., Hamilton, D. P., and Taylor, P. A. (2008). Cooling lakes while the world warms : Effects of forest regrowth and increased dissolved organic matter on the thermal regime of a temperate, urban lake. *Limnology and Oceanography*, 53(1) :404–410.
- Anis, A. and Moum, J. N. (1995). Surface Wave–Turbulence Interactions. Scaling  $\epsilon(z)$  near the Sea Surface. *Journal of Physical Oceanography*, 25(9) :2025–2045.
- Arsenev, S. and Felzenbaum, A. I. (1977). Integral model of the active ocean layer. *Academy of Sciences, USSR, Izvestiya, Atmospheric and Oceanic Physics*, (13) :1034–1043.
- Bastviken, D., Cole, J., Pace, M., and Tranvik, L. (2004). Methane emissions from lakes : Dependence of lake characteristics, two regional assessments, and a global estimate : Lake Methane Emissions. *Global Biogeochemical Cycles*, 18(4).
- Biancamaria, S., Lettenmaier, D. P., and Pavelsky, T. M. (2016). The SWOT Mission and Its Capabilities for Land Hydrology. *Surveys in Geophysics*, 37(2) :307–337.
- Blumberg, A. F. and Mellor, G. L. (1987). A description of a three-dimensional coastal ocean circulation model. In Heaps, N. S., editor, *Coastal and Estuarine Sciences*, volume 4, pages 1–16. American Geophysical Union, Washington, D. C.
- Boegman, L. (2009). Currents in Stratified Water Bodies 2 : Internal Waves. In *Encyclopedia of Inland Waters*, pages 539–558. Elsevier.
- Bonan, G. B. (1995). Sensitivity of a GCM Simulation to Inclusion of Inland Water Surfaces. *Journal of Climate*, 8(11) :2691–2704.

- Bowling, L. C. and Lettenmaier, D. P. (2010). Modeling the Effects of Lakes and Wetlands on the Water Balance of Arctic Environments. *Journal of Hydrometeorology*, 11(2) :276–295.
- Bradley, E. F., Coppin, P. A., and Godfrey, J. S. (1991). Measurements of sensible and latent heat flux in the western equatorial Pacific Ocean. *Journal of Geophysical Research*, 96(S01) :3375.
- Brooks, J. R., Gibson, J. J., Birks, S. J., Weber, M. H., Rodecap, K. D., and Stoddard, J. L. (2014). Stable isotope estimates of evaporation : inflow and water residence time for lakes across the United States as a tool for national lake water quality assessments. *Limnology and Oceanography*, 59(6) :2150–2165.
- Buchard, H. (2001). Simulating the Wave-Enhanced Layer under Breaking Surface Waves with Two-Equation Turbulence Models. *Journal of Physical Oceanography*, 31(11) :3133–3145.
- Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F. (1971). Flux-Profile Relationships in the Atmospheric Surface Layer. *Journal of the Atmospheric Science*, 28(2) :181–189.
- Charnock, H. (1955). Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, 81(350) :639–640.
- Cherkauer, K. A., Bowling, L. C., and Lettenmaier, D. P. (2003). Variable infiltration capacity cold land process model updates. *Global and Planetary Change*, 38(1-2) :151–159.
- Choulga, M., Kourzeneva, E., Balsamo, G., Boussetta, S., and Wedi, N. (2019). Upgraded global mapping information for earth system modelling : an application to surface water depth at the ECMWF. *Hydrology and Earth System Sciences*, 23(10) :4051–4076.
- Chuang, H.-Y. and Sousounis, P. J. (2003). The Impact of the Prevailing Synoptic Situation on the Lake-Aggregate Effect. *Monthly Weather Review*, 131(5) :990–1010.
- 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. (2007). Plumbing the Global Carbon Cycle : Integrating Inland Waters into the Terrestrial Carbon Budget. *Ecosystems*, 10(1) :172–185.
- Cretaux, J.-F., Nielsen, K., Frappart, F., Papa, F., Calmant, S., and Benveniste, J. (2017). Hydrological applications of satellite altimetry rivers, lakes, man-made reservoirs, inundated areas. In *Satellite altimetry over oceans and land surfaces*.

- In *Satellite Altimetry over Oceans and Land Surfaces*, pages 459–504. CRC Press.
- Crétaux, J.-F., Arsen, A., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, R., Cazenave, A., and Maison-grande, P. (2011). SOLS : A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Advances in Space Research*, 47(9) :1497–1507.
- de Rosnay, P. (2003). Integrated parameterization of irrigation in the land surface model ORCHIDEE. Validation over Indian Peninsula. *Geophysical Research Letters*, 30(19) :1986.
- Deardorff, J. W. (1970). Preliminary Results from Numerical Integrations of the Unstable Planetary Boundary Layer. *Journal of the Atmospheric Sciences*, 27(8) :1209–1211.
- Deardorff, J. W., Willis, G. E., and Stockton, B. H. (1980). Laboratory studies of the entrainment zone of a convectively mixed layer. *Journal of Fluid Mechanics*, 100(01) :41.
- Delsol, F., Miyakoda, K., and Clarke, R. H. (1971). Parameterized processes in the surface boundary layer of an atmospheric circulation model. *Quarterly Journal of the Royal Meteorological Society*, 97(412) :181–208.
- DelSontro, T., Boutet, L., St-Pierre, A., del Giorgio, P. A., and Prairie, Y. T. (2016). Methane ebullition and diffusion from northern ponds and lakes regulated by the interaction between temperature and system productivity : Productivity regulates methane lake flux. *Limnology and Oceanography*, 61(S1) :S62–S77.
- Dokulil, M. T. (2014). Predicting summer surface water temperatures for large Austrian lakes in 2050 under climate change scenarios. *Hydrobiologia*, 731(1) :19–29.
- Dokulil, M. T., Jagsch, A., George, G. D., Anneville, O., Jankowski, T., Wahl, B., Lenhart, B., Blenckner, T., and Teubner, K. (2006). Twenty years of spatially coherent deepwater warming in lakes across Europe related to the North Atlantic Oscillation. *Limnology and Oceanography*, 51(6) :2787–2793.
- 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. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*, 51(5) :2388–2397.

- Ducoudré, N. I., Laval, K., and Perrier, A. (1993). SECHIBA, a new set of parameterizations of the hydrologic exchanges at the land-atmosphere interface within the LMD atmospheric general circulation model. *Journal of Climate*, 6 :248–273.
- Dupont, F., Chittibabu, P., Fortin, V., Rao, Y. R., and Lu, Y. (2012). Assessment of a NEMO-based hydrodynamic modelling system for the Great Lakes. *Water Quality Research Journal*, 47(3-4) :198–214.
- Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P., Mironov, D., and Schar, C. (2010). An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal Environment Research*, 15(2) :100–112.
- ESA (2017). Land cover CCI product user guide version 2. Technical report, ESA.
- Faroux, S., Kaptué Tchuenté, A. T., Roujean, J.-L., Masson, V., Martin, E., and Le Moigne, P. (2013). ECOCLIMAP-II/Europe : a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. *Geoscientific Model Development*, 6(2) :563–582.
- Filyushkin, B. N. and Miropolsky, Y. Z. (1981). Seasonal variability of the upper thermocline and self-similarity of temperature profiles. *Okeanologia*, 21 :416–424.
- Fischer, H. B., List, J. E., Koh, C. R., Imberger, J., and Brooks, N. H. (1979). *Mixing in inland and coastal waters*. Academic press.
- Fjortoft, R., Gaudin, J.-M., Pourthie, N., Lalaurie, J.-C., Mallet, A., Nouvel, J.-F., Martinot-Lagarde, J., Oriot, H., Borderies, P., Ruiz, C., and Daniel, S. (2014). KaRIn on SWOT : Characteristics of Near-Nadir Ka-Band Interferometric SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(4) :2172–2185.
- Friedlingstein, P., Houghton, R. A., Marland, G., Hackler, J., Boden, T. A., Conway, T. J., Canadell, J. G., Raupach, M. R., Ciais, P., and Le Quéré, C. (2010). Update on CO<sub>2</sub> emissions. *Nature Geoscience*, 3(12) :811–812.
- Froehlich, K., Gonfiantini, R., and Rozanski, K. (2005). Isotopes in Lake Studies : A Historical Perspective. In Aggarwal, P. K., Gat, J. R., and Froehlich, K. F., editors, *Isotopes in the Water Cycle*, pages 139–150. Springer-Verlag, Berlin/Heidelberg.
- Frohlich, C. and Lean, J. (2004). Solar radiative output and its variability : evidence and mechanisms. *The Astronomy and Astrophysics Review*, 12(4) :273–320.

- Fu, L.-L., Alsdorf, D., Morrow, R., Rodriguez, E., and Mognard, N. (2012). SWOT : the Surface Water and Ocean Topography Mission : wide- swath altimetric elevation on Earth. Technical Report, NASA/JPL.
- 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. (2019). The DOE E3SM Coupled Model Version 1 : Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems*, 11(7) :2089–2129.
- Golosov, S., Zverev, I., Shipunova, E., and Terzhevik, A. (2018). Modified parameterization of the vertical water temperature profile in the FLake model. *Tellus A : Dynamic Meteorology and Oceanography*, 70(1) :1–7.
- Goudsmit, G.-H., Burchard, H., Peeters, F., and Wüest, A. (2002). Application of k- $\epsilon$  turbulence models to enclosed basins : The role of internal seiches : APPLICATION OF k- $\epsilon$  TURBULENCE MODELS. *Journal of Geophysical Research : Oceans*, 107(C12) :23–1–23–13.
- Goyette, S., McFarlane, N., and Flato, G. M. (2000). Application of the Canadian regional climate model to the Laurentian great lakes region : Implementation of a lake model. *Atmosphere-Ocean*, 38(3) :481–503.
- Gronewold, A. D., Smith, J. P., Read, L. K., and Crooks, J. L. (2020). Reconciling the water balance of large lake systems. *Advances in Water Resources*, 137 :103505.
- Gueymard, C. A. (2018). A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. *Solar Energy*, 168 :2–9.
- Guinaldo, T., Munier, S., Le Moigne, P., Boone, A., Decharme, B., Choulga, M., and Leroux, D. J. (2021). Parametrization of a lake water dynamics model

- MLake in the ISBA-CTRIP land surface system (SURFEX v8.1). *Geoscientific Model Development*, 14(3) :1309–1344.
- Gurvan, M., 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. (2017). NEMO ocean engine.
- He, B., Bao, Q., Wang, X., Zhou, L., Wu, X., Liu, Y., Wu, G., Chen, K., He, S., Hu, W., Li, J., Li, J., Nian, G., Wang, L., Yang, J., Zhang, M., and Zhang, X. (2019). CAS FGOALS-f3-L Model Datasets for CMIP6 Historical Atmospheric Model Intercomparison Project Simulation. *Advances in Atmospheric Sciences*, 36(8) :771–778.
- Heathcote, A. J., del Giorgio, P. A., and Prairie, Y. T. (2015). Predicting bathymetric features of lakes from the topography of their surrounding landscape. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(5) :643–650.
- 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. (2015). Effects of water clarity on lake stratification and lake-atmosphere heat exchange. *Journal of Geophysical Research : Atmospheres*, 120(15) :7412–7428.
- Henderson-Sellers, B. (1985). New formulation of eddy diffusion thermocline models. *Applied Mathematical Modelling*, 9(6) :441–446.
- Hipsey, M. R., Bruce, L. C., Boon, C., Busch, B., Carey, C. C., Hamilton, D. P., Hanson, P. C., Read, J. S., de Sousa, E., Weber, M., and Winslow, L. A. (2019). A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). *Geoscientific Model Development*, 12(1) :473–523.
- Hodges, B. R., Imberger, J., Saggio, A., and Winters, K. B. (2000). Modeling basin-scale internal waves in a stratified lake. *Limnology and Oceanography*, 45(7) :1603–1620.
- Hondzo, M. and Stefan, H. G. (1993). Lake Water Temperature Simulation Model. *Journal of Hydraulic Engineering*, 119(11) :1251–1273.
- Hostetler, S. W. and Bartlein, P. J. (1990). Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur Lake, Oregon. *Water Resources Research*, 26(10) :2603–2612.

- Hutchinson, G. E. (1967). *A treatise on limnology, introduction to lake biology and the limnoplankton*, volume 2. Wiley edition.
- Hutchinson, G. E. and Loffler, H. (1956). THE THERMAL CLASSIFICATION OF LAKES. *Proceedings of the National Academy of Sciences*, 42(2) :84–86.
- Huziy, O. and Sushama, L. (2017). Impact of lake–river connectivity and interflow on the Canadian RCM simulated regional climate and hydrology for Northeast Canada. *Climate Dynamics*, 48(3-4) :709–725.
- Imberger, J. (1985). The diurnal mixed layer<sup>1</sup> : The diurnal mixed layer. *Limnology and Oceanography*, 30(4) :737–770.
- Intergovernmental Panel on Climate Change, editor (2014). *Climate Change 2013 - The Physical Science Basis : Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.
- Iooss, B. and Lemaitre, P. (2015). A Review on Global Sensitivity Analysis Methods. In Dellino, G. and Meloni, C., editors, *Uncertainty Management in Simulation-Optimization of Complex Systems*, volume 59, pages 101–122. Springer US, Boston, MA.
- Joehnk, K. and Umlauf, L. (2001). Modelling the metalimnetic oxygen minimum in a medium sized alpine lake. *Ecological Modelling*, 136(1) :67–80.
- Kitaigorodskii, S. A. and Miropolsky, Y. Z. (1970). On the theory of the open ocean active layer. 6(2) :178.
- Kourzeneva, e. (2010). external data for lake parameterization in numerical Weather Prediction and climate modeling.
- Krinner, G. (2003). Impact of lakes and wetlands on boreal climate. *Journal of Geophysical Research*, 108(D16).
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system : DVGM FOR COUPLED CLIMATE STUDIES. *Global Biogeochemical Cycles*, 19(1).
- Lamarche, C., Santoro, M., Bontemps, S., d'Andrimont, R., Radoux, J., Giustarini, L., Brockmann, C., Wevers, J., Defourny, P., and Arino, O. (2017). Compilation and Validation of SAR and Optical Data Products for a Complete and Global Map of Inland/Ocean Water Tailored to the Climate Modeling Community. *Remote Sensing*, 9(1) :36.



- Larson, G. and Schaetzl, R. (2001). Origin and Evolution of the Great Lakes. *Journal of Great Lakes Research*, 27(4) :518–546.
- Laval, K., Sadourny, R., and Serafini, Y. (1981). Land surface processes in a simplified general circulation model. *Geophysical & Astrophysical Fluid Dynamics*, 17(1) :129–150.
- Le Moigne, P., Colin, J., and Decharme, B. (2016). Impact of lake surface temperatures simulated by the FLake scheme in the CNRM-CM5 climate model. *Tellus A : Dynamic Meteorology and Oceanography*, 68(1) :31274.
- Leavitt, P. R., Fritz, S. C., Anderson, N. J., Baker, P. A., Blenckner, T., Bunting, L., Catalan, J., Conley, D. J., Hobbs, W. O., Jeppesen, E., Korhola, A., McGowan, S., RÜhland, K., Rusak, J. A., Simpson, G. L., Solovieva, N., and Werne, J. (2009). Paleolimnological evidence of the effects on lakes of energy and mass transfer from climate and humans. *Limnology and Oceanography*, 54(6part2) :2330–2348.
- Lehner, B. and Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296(1-4) :1–22.
- Lenters, J. D., Kratz, T. K., and Bowser, C. J. (2005). Effects of climate variability on lake evaporation : Results from a long-term energy budget study of Sparkling Lake, northern Wisconsin (USA). *Journal of Hydrology*, 308(1-4) :168–195.
- Leon, L., Lam, D., Schertzer, W., Swayne, D., and Imberger, J. (2007). Towards coupling a 3D hydrodynamic lake model with the Canadian Regional Climate Model : Simulation on Great Slave Lake. *Environmental Modelling & Software*, 22(6) :787–796.
- Lesser, G., Roelvink, J., van Kester, J., and Stelling, G. (2004). Development and validation of a three-dimensional morphological model. *Coastal Engineering*, 51(8-9) :883–915.
- Lewis Jr., W. M. (1983). A Revised Classification of Lakes Based on Mixing. *Canadian Journal of Fisheries and Aquatic Sciences*, 40(10) :1779–1787.
- Li, W., Zhang, Y., Shi, X., Zhou, W., Huang, A., Mu, M., Qiu, B., and Ji, J. (2019). Development of Land Surface Model BCC\_avim2.0 and Its Preliminary Performance in LS3MIP/CMIP6. *Journal of Meteorological Research*, 33(5) :851–869.
- Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99(D7) :14415.

- Ljungemyr, P., Gustafsson, N., and Omstedt, A. (1996). Parameterization of lake thermodynamics in a high-resolution weather forecasting model. *Tellus A : Dynamic Meteorology and Oceanography*, 48(5) :608–621.
- Long, Z., Perrie, W., Gyakum, J., Caya, D., and Laprise, R. (2007). Northern Lake Impacts on Local Seasonal Climate. *Journal of Hydrometeorology*, 8(4) :881–896.
- Louis, J.-F. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, 17(2) :187–202.
- 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. (2020). Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model. *Journal of Advances in Modeling Earth Systems*, 12(4).
- MacCallum, S. N. and Merchant, C. J. (2012). Surface water temperature observations of large lakes by optimal estimation. *Canadian Journal of Remote Sensing*, 38(1) :25–45.
- MacKay, M. D. (2012). A Process-Oriented Small Lake Scheme for Coupled Climate Modeling Applications. *Journal of Hydrometeorology*, 13(6) :1911–1924.
- MacKay, M. D., Versegny, D. L., Fortin, V., and Rennie, M. D. (2017). Wintertime Simulations of a Boreal Lake with the Canadian Small Lake Model. *Journal of Hydrometeorology*, 18(8) :2143–2160.
- Malkki, P. and Tamsalu, R. E. (1985). *Physical Features of the Baltic Sea*. 252. Finnish marine research, edition.
- Martynov, A., Sushama, L., and Laprise, R. (2010). Simulation of temperate freezing lakes by one-dimensional lake models : performance assessment for interactive coupling with regional climate models. *Boreal Environment Research*, 15(1) :143–164.
- Martynov, A., Sushama, L., Laprise, R., Winger, K., and Dugas, B. (2012). 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(1) :16226.
- Mats, V. D., Shcherbakov, D. Y., and Efimova, I. M. (2011). Late Cretaceous-Cenozoic history of the Lake Baikal depression and formation of its unique biodiversity. *Stratigraphy and Geological Correlation*, 19(4) :404–423.

- Messenger, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7(1) :13603.
- Millero, F. J., Poisson, A., Chen, C. T., Bradshaw, A. L., and Schleicher, K. (1981). Background papers and supporting data on the international equation of state of seawater 1980. Technical report, UNESCO.
- 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. (2014). An Enhanced Model of Land Water and Energy for Global Hydrologic and Earth-System Studies. *Journal of Hydrometeorology*, 15(5) :1739–1761.
- Mironov, D. (2008). Parameterization of lakes in numerical weather prediction. Description of a lake model. Technical Report 11, Deutscher Wetterdienst.
- Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., and Terzhevik, A. (2010). Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environment Research*, 15(2) :218–230.
- Molod, A. (2002). A global assessment of the mosaic approach to modeling land surface heterogeneity. *Journal of Geophysical Research*, 107(D14) :4217.
- Monsi, M. and Saeki, T. (1953). Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion. *Japanese Journal of Botany*, 14 :22–52.
- Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., and Fita, L. (2018). ORCHIDEE-ROUTING : revising the river routing scheme using a high-resolution hydrological database. *Geoscientific Model Development*, 11(12) :4965–4985.
- Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F. (2001). Predicting the Discharge of Global Rivers. *Journal of Climate*, 14(15) :3307–3323.
- North, R. P., North, R. L., Livingstone, D. M., Köster, O., and Kipfer, R. (2014). Long-term changes in hypoxia and soluble reactive phosphorus in the hypolimnion of a large temperate lake : consequences of a climate regime shift. *Global Change Biology*, 20(3) :811–823.
- 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.

- (2013). Technical description of version 4.5 of the Community Land Model (CLM). Technical report, UCAR/NCAR.
- O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., Schneider, P., Lenters, J. D., McIntyre, P. B., Kraemer, B. M., Weyhenmeyer, G. A., Straile, D., Dong, B., Adrian, R., Allan, M. G., Anneville, O., Arvola, L., Austin, J., Bailey, J. L., Baron, J. S., Brookes, J. D., Eyto, E., Dokulil, M. T., Hamilton, D. P., Havens, K., Hetherington, A. L., Higgins, S. N., Hook, S., Izmet'eva, L. R., Joehnk, K. D., Kangur, K., Kasprzak, P., Kumagai, M., Kuusisto, E., Leshkevich, G., Livingstone, D. M., MacIntyre, S., May, L., Melack, J. M., Mueller-Navarra, D. C., Naumenko, M., Noges, P., Noges, T., North, R. P., Plisnier, P., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L. G., Rusak, J. A., Salmaso, N., Samal, N. R., Schindler, D. E., Schladow, S. G., Schmid, M., Schmidt, S. R., Silow, E., Soylu, M. E., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Williamson, C. E., and Zhang, G. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, 42(24).
- Ottlé, C., Bernus, A., Verbeke, T., Pétrus, K., Yin, Z., Biancamaria, S., Jost, A., Desroches, D., Pottier, C., Perrin, C., de Lavenne, A., Flipo, N., and Rivière, A. (2020). Characterization of SWOT Water Level Errors on Seine Reservoirs and La Bassée Gravel Pits : Impacts on Water Surface Energy Budget Modeling. *Remote Sensing*, 12(18) :2911.
- Park, S., Shin, J., Kim, S., Oh, E., and Kim, Y. (2019). Global Climate Simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON). *Journal of Climate*, 32(10) :2917–2949.
- Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633) :418–422.
- Petermann, E., Gibson, J. J., Knöller, K., Pannier, T., Weiß, H., and Schubert, M. (2018). Determination of groundwater discharge rates and water residence time of groundwater-fed lakes by stable isotopes of water ( $^{18}\text{O}$ ,  $^2\text{H}$ ) and radon ( $^{222}\text{Rn}$ ) mass balances. *Hydrological Processes*, 32(6) :805–816.
- Piccolroaz, S., Toffolon, M., and Majone, B. (2013). A simple lumped model to convert air temperature into surface water temperature in lakes. *Hydrology and Earth System Sciences*, 17(8) :3323–3338.
- Piccolroaz, S., Toffolon, M., and Majone, B. (2015). The role of stratification on lakes' thermal response : The case of Lake Superior. *Water Resources Research*, 51(10) :7878–7894.

- Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H., Laaksonen, A., and Kaurola, J. (2018). 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(4) :1321–1342.
- Pitman, A. (1991). A simple parameterization of sub-grid scale open water for climate models. *Climate Dynamics*, 6(2) :99–112.
- Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., and Zimmermann, N. E. (2011). Plant functional type mapping for earth system models. *Geoscientific Model Development*, 4(4) :993–1010.
- Rasilo, T., Prairie, Y. T., and Giorgio, P. A. (2015). Large-scale patterns in summer diffusive  $\text{CH}_4$  fluxes across boreal lakes, and contribution to diffusive C emissions. *Global Change Biology*, 21(3) :1124–1139.
- Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjaeld, S. (2021). JSBACH 3 - The land component of the MPI Earth System Model : documentation of version 3.2. page 4990986.
- Richardson, D., Melles, S., Pilla, R., Hetherington, A., Knoll, L., Williamson, C., Kraemer, B., Jackson, J., Long, E., Moore, K., Rudstam, L., Rusak, J., Saros, J., Sharma, S., Strock, K., Weathers, K., and Wigdahl-Perry, C. (2017). Transparency, Geomorphology and Mixing Regime Explain Variability in Trends in Lake Temperature and Stratification across Northeastern North America (1975–2014). *Water*, 9(6) :442.
- Robinson, N., Regetz, J., and Guralnick, R. P. (2014). EarthEnv-DEM90 : A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87 :57–67.
- Rong, X., Li, J., Chen, H., Xin, Y., Su, J., Hua, L., Zhou, T., Qi, Y., Zhang, Z., Zhang, G., and Li, J. (2018). The CAMS Climate System Model and a Basic Evaluation of Its Climatology and Climate Variability Simulation. *Journal of Meteorological Research*, 32(6) :839–861.
- Rooney, G. and Jones, I. (2010). Coupling the 1-D lake model FLake to the community land-surface model JULES. *Boreal Environment Research*, 15(5) :501–512.
- Sachse, R., Petzoldt, T., Blumstock, M., Moreira, S., Pätzig, M., Rücker, J., Janse, J. H., Mooij, W. M., and Hilt, S. (2014). Extending one-dimensional models for

- deep lakes to simulate the impact of submerged macrophytes on water quality. *Environmental Modelling & Software*, 61 :410–423.
- Salerno, F., Gambelli, S., Viviano, G., Thakuri, S., Guyennon, N., D'Agata, C., Diolaiuti, G., Smiraglia, C., Stefani, F., Bocchiola, D., and Tartari, G. (2014). High alpine ponds shift upwards as average temperatures increase : A case study of the Ortles–Cevedale mountain group (Southern Alps, Italy) over the last 50years. *Global and Planetary Change*, 120 :81–91.
- Salgado, R. and Le Moigne, P. (2010). Coupling of the Flake model to the surfex externalized surface model. *Boreal Environ Research Publishing Board*, 15 :231–244.
- Saloranta, T. M. and Andersen, T. (2007). MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecological Modelling*, 207(1) :45–60.
- Seekell, D. A., Pace, M. L., Tranvik, L. J., and Verpoorter, C. (2013). A fractal-based approach to lake size-distributions : LAKE SIZE-DISTRIBUTIONS. *Geophysical Research Letters*, 40(3) :517–521.
- 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. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, 13(12) :6165–6200.
- Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahann, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M. (2019). UKESM1 : Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, 11(12) :4513–4558.
- Semmler, T., Cheng, B., Yang, Y., and Rontu, L. (2012). Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models. *Tellus A : Dynamic Meteorology and Oceanography*, 64(1) :17339.

- Sheng, Y., Song, C., Wang, J., Lyons, E. A., Knox, B. R., Cox, J. S., and Gao, F. (2016). Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sensing of Environment*, 185 :129–141.
- Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model : LPJ DYNAMIC GLOBAL VEGETATION MODEL. *Global Change Biology*, 9(2) :161–185.
- Song, Y., Semazzi, F. H. M., Xie, L., and Ogallo, L. J. (2004). A coupled regional climate model for the Lake Victoria basin of East Africa : COUPLED REGIONAL CLIMATE MODEL FOR LAKE VICTORIA BASIN. *International Journal of Climatology*, 24(1) :57–75.
- Spence, C. (2006). Hydrological processes and streamflow in a lake dominated watercourse. *Hydrological Processes*, 20(17) :3665–3681.
- Spigel, R. H. and Imberger, J. (1980). The classification of Mixed-Layer Dynamics of Lakes of Small to Medium Size. *Journal of Physical Oceanography*, 10(7) :1104–1121.
- Stacey, M. W. (1999). Simulation of the Wind-Forced Near-Surface Circulation in Knight Inlet : A Parameterization of the Roughness Length. *Journal of Physical Oceanography*, 29(6) :1363–1367.
- Stepanenko, V., Mammarella, I., Ojala, A., Miettinen, H., Lykosov, V., and Vesala, T. (2016). LAKE 2.0 : a model for temperature, methane, carbon dioxide and oxygen dynamics in lakes. *Geoscientific Model Development*, 9(5) :1977–2006.
- Stepanenko, V. M., Martynov, A., Jöhnk, K. D., Subin, Z. M., Perroud, M., Fang, X., Beyrich, F., Mironov, D., and Goyette, S. (2013). A one-dimensional model intercomparison study of thermal regime of a shallow, turbid midlatitude lake. *Geoscientific Model Development*, 6(4) :1337–1352.
- Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., L'Ecuyer, T., Stackhouse, P. W., Lebsock, M., and Andrews, T. (2012). An update on Earth's energy balance in light of the latest global observations. *Nature Geoscience*, 5(10) :691–696.
- Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S. (2003). Global distribution of C<sub>3</sub> and C<sub>4</sub> vegetation : Carbon cycle implications : C<sub>4</sub> PLANTS AND CARBON CYCLE. *Global Biogeochemical Cycles*, 17(1) :6–16–14.

- Subin, Z. M., Riley, W. J., and Mironov, D. (2012). An improved lake model for climate simulations : Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems*, 4 :M02001.
- Tan, Z., Zhuang, Q., Shurpali, N. J., Marushchak, M. E., Biasi, C., Eugster, W., and Walter Anthony, K. (2017). Modeling  $\text{CO}_2$  emissions from Arctic lakes : Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 9(5) :2190–2213.
- Tan, Z., Zhuang, Q., and Walter Anthony, K. (2015). Modeling methane emissions from arctic lakes : Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 7(2) :459–483.
- Terray, E. A., Donelan, M. A., Agrawal, Y. C., Drennan, W. M., Kahma, K. K., Williams, A. J., Hwang, P. A., and Kitaigorodskii (1996). Estimates of Kinetic Energy Dissipation under Breaking Waves. *Journal of Physical Oceanography*, 26(5) :792–807.
- Thiery, W., Davin, E. L., Panitz, H.-J., Demuzere, M., Lhermitte, S., and van Lipzig, N. (2015). The Impact of the African Great Lakes on the Regional Climate. *Journal of Climate*, 28(10) :4061–4085.
- Thompson, R., Ventura, M., and Camarero, L. (2009). On the climate and weather of mountain and sub-arctic lakes in Europe and their susceptibility to future climate change. *Freshwater Biology*, 54(12) :2433–2451.
- Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., Dillon, P., Finlay, K., Fortino, K., Knoll, L. B., Kortelainen, P. L., Kutser, T., Larsen, S., Laurion, I., Leech, D. M., McCallister, S. L., McKnight, D. M., Melack, J. M., Overholt, E., Porter, J. A., Prairie, Y., Renwick, W. H., Roland, F., Sherman, B. S., Schindler, D. W., Sobek, S., Tremblay, A., Vanni, M. J., Verschoor, A. M., von Wachenfeldt, E., and Weyhenmeyer, G. A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, 54(6part2) :2298–2314.
- Tsuang, B.-J. and Arpe, K. (2001). Lake parameterization for climate models. Technical Report 316, Max-Planck-Institut für Meteorologie.
- Van de Walle, J., Thiery, W., Brousse, O., Souverijns, N., Demuzere, M., and van Lipzig, N. P. M. (2020). A convection-permitting model for the Lake Victoria Basin : evaluation and insight into the mesoscale versus synoptic atmospheric dynamics. *Climate Dynamics*, 54(3-4) :1779–1799.



- Vanderkelen, I., Lipzig, N. P. M., Lawrence, D. M., Droppers, B., Golub, M., Gosling, S. N., Janssen, A. B. G., Marcé, R., Schmied, H. M., Perroud, M., Pierson, D., Pokhrel, Y., Satoh, Y., Schewe, J., Seneviratne, S. I., Stepanenko, V. M., Tan, Z., Woolway, R. I., and Thiery, W. (2020). Global Heat Uptake by Inland Waters. *Geophysical Research Letters*, 47(12).
- Verpoorter, C., Kutser, T., Seekell, D. A., and Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18) :6396–6402.
- Viovy, N. (1996). Interannuality and CO<sub>2</sub> sensitivity of the SECHIBA-BGC coupled SVAT-BGC model. *Physics and Chemistry of the Earth*, 21(5-6) :489–497.
- Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnavé, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F. (2013). The CNRM-CM5.1 global climate model : description and basic evaluation. *Climate Dynamics*, 40(9-10) :2091–2121.
- Vuichard, N., Messina, P., Luysaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., and Peylin, P. (2019). 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(11) :4751–4779.
- Wagner, C. and Adrian, R. (2009). Cyanobacteria dominance : Quantifying the effects of climate change. *Limnology and Oceanography*, 54(6part2) :2460–2468.
- Wang, Z.-T., Chen, T.-Y., Liu, S.-W., and Lai, Z.-P. (2016). Aeolian origin of interdune lakes in the Badain Jaran Desert, China. *Arabian Journal of Geosciences*, 9(3) :190.
- Weckström, K., Weckström, J., Huber, K., Kamenik, C., Schmidt, R., Salvenmoser, W., Rieradevall, M., Weisse, T., Psenner, R., and Kurmayer, R. (2016). Impacts of Climate Warming on Alpine Lake Biota Over the Past Decade. *Arctic, Antarctic, and Alpine Research*, 48(2) :361–376.
- West, W. E., Creamer, K. P., and Jones, S. E. (2016). Productivity and depth regulate lake contributions to atmospheric methane : Lake productivity fuels methane emissions. *Limnology and Oceanography*, 61(S1) :S51–S61.
- Woolway, R. I., Dokulil, M. T., Marszelewski, W., Schmid, M., Bouffard, D., and Merchant, C. J. (2017). Warming of Central European lakes and their response to the 1980s climate regime shift. *Climatic Change*, 142(3-4) :505–520.

- Woolway, R. I. and Merchant, C. J. (2019). Worldwide alteration of lake mixing regimes in response to climate change. *Nature Geoscience*, 12(4) :271–276.
- Wright, D. M., Posselt, D. J., and Steiner, A. L. (2013). Sensitivity of Lake-Effect Snowfall to Lake Ice Cover and Temperature in the Great Lakes Region. *Monthly Weather Review*, 141(2) :670–689.
- Wu, J. (1994). The sea surface is aerodynamically rough even under light winds. *Boundary-Layer Meteorology*, 69(1-2) :149–158.
- Wüest, A. and Lorke, A. (2003). SMALL-SCALE HYDRODYNAMICS IN LAKES. *Annual Review of Fluid Mechanics*, 35(1) :373–412.
- Wüest, A., Ravens, T. M., Granin, N. G., Kocsis, O., Schurter, M., and Sturm, M. (2005). Cold intrusions in Lake Baikal : Direct observational evidence for deep-water renewal. *Limnology and Oceanography*, 50(1) :184–196.
- Yankova, Y., Neuenschwander, S., Köster, O., and Posch, T. (2017). Abrupt stop of deep water turnover with lake warming : Drastic consequences for algal primary producers. *Scientific Reports*, 7(1) :13770.
- Yao, X., Liu, S., Han, L., Sun, M., and Zhao, L. (2018). Definition and classification system of glacial lake for inventory and hazards study. *Journal of Geographical Sciences*, 28(2) :193–205.
- Yeates, P. and Imberger, J. (2003). Pseudo two-dimensional simulations of internal and boundary fluxes in stratified lakes and reservoirs. *International Journal of River Basin Management*, 1(4) :297–319.
- Yelland, M. and Taylor, P. K. (1996). Wind Stress Measurements from the Open Ocean. *Journal of Physical Oceanography*, 26(4) :541–558.
- Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C., Yang, K., Piao, S., Sheng, Y., Yi, S., Li, J., O'Reilly, C. M., Qi, S., Shen, S. S., Zhang, H., and Jia, Y. (2019). Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes. *Remote Sensing of Environment*, 221 :386–404.
- Zhong, Y., Notaro, M., Vavrus, S. J., and Foster, M. J. (2016). Recent accelerated warming of the Laurentian Great Lakes : Physical drivers. *Limnology and Oceanography*, 61(5) :1762–1786.
- Zolfaghari, K., Duguay, C. R., and Kheyrollah Pour, H. (2017). 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.

Zou, Z., Xiao, X., Dong, J., Qin, Y., Doughty, R. B., Menarguez, M. A., Zhang, G., and Wang, J. (2018). Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. *Proceedings of the National Academy of Sciences*, 115(15) :3810–3815.