Using automatic calibration to improve the physics behind complex numerical models: An example from a 3D lake model using ~~Delft3d~~ Delft3D (v6.02.10) and DYNO-PODS (v1.0)

Marina Amadori[1,2], Abolfazl Irani Rahaghi[3,4], Damien Bouffard[3,5], and Marco Toffolon[2]

[1]Institute for Electromagnetic Sensing of the Environment, National Research Council, Milan, Italy
[2]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy
[3]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Surface Waters – Research and Management, Switzerland
[4]Department of Geography, University of Zurich, 8057, Zurich, Switzerland
[5]Faculty of Geosciences and Environment, Institute of Earth Surface Dynamics, University of Lausanne, Geopolis, Mouline, 1015, Lausanne, Switzerland

Correspondence: Marina Amadori (marina.amadori@unitn.it)

Abstract. Models are simplified descriptions of reality and are intrinsically limited by the assumptions that have been introduced in their formulation. With the development of automatic calibration toolboxes, finding optimal parameters that suit the environmental system has become more convenient. Here, we explore how optimization toolboxes can be applied innovatively to uncover flaws in the physical formulations of models. We illustrate this approach by evaluating the effect of simplifications embedded in the formulation of a widely used hydro-thermodynamic model. We calibrate a Delft3D model based on temperature profiles for a case study, Lake Morat (Switzerland), through the ~~optimization tool DYNO-PODS. Results show that~~ DYNO-PODS optimization tool. The results show that higher values of the light extinction coefficient can compensate for neglecting the fraction β of ~~shortwave~~ short-wave radiation absorbed at the ~~water surface can be compensated by higher values of the light extinction coefficient~~ surface of the water. This leads to unrealistic values of the ~~latter parameter~~ light extinction coefficient, as the optimization pushes ~~the coefficient~~ its value towards the limit of no transparency, consistent with the need to reproduce a significant absorption at the surface. ~~While it is well-known~~ Although it is well known that β is significantly larger than zero, its absence ~~in~~ from the model was never noticed as critical. ~~The extensive use of automatic calibration tools may offer similar outcomes in other applications~~ Automatic calibration tools provide valuable diagnostic insights into the physical robustness of models, enabling more precise evaluation of their structural integrity and performance.

1 Introduction

Numerical models serve as powerful tools capable of embracing the complexity of intricate environmental dynamics. In many branches of environmental science, such as meteorology, climatology, hydrology, oceanography, and limnology, ~~thermo-hydrodynamic~~ hydro-thermodynamic models have become standard tools to simulate and understand specific physical processes. These models are rooted on the numerical solution of physical first principles, e.g. the mass, momentum, and heat equations. Although

first principles are well established, modeling the physics of fluid systems remains complicated for two reasons. First, the grid size limits the range of applicability of models based on first principles (e.g., Direct Numerical Simulations, DNS) and poses the challenge to properly parameterize associated sub-grid scale processes. This implies that models that were developed for a specific environment, where the parametrization was adequate, may not be optimized for other contexts characterized by different spatial scales. Second, imposing the boundary conditions is a complicated task as the forcing acting at the boundaries of the computational domain is often only partially known. For instance, the estimation of the surface heat fluxes is based on parametrizations that have to cope with both the complex physics and the uncertainties associated with the forcing. Contrary to the first principles solved by the model, such boundary conditions often lack universality. In the specific case of models widely adopted to simulate three-dimensional lake dynamics, many were originally developed for marine, riverine or estuarine environments. This is the case of, e.g., Delft3D (Lesser et al., 2004), MITgcm (Marshall et al., 1997), ROMS (Shchepetkin and McWilliams, 2005), FVCOM (Chen et al., 2003), NEMO (Madec et al., 2023), POM (Blumberg and Mellor, 1987), among a few others. Therefore, some processes that are crucial in lakes, but not for the original environment, may be correctly reproduced. Scientists thus must critically evaluate model performances, as even some of the most used models may have unexpected flaws.

The first step of any model setup and performance evaluation is calibration. Tools for the automatic calibration of the model parameters have recently been introduced, enabling a broader and more efficient search for the optimal parameters. Analyses of how these tools can provide support and foster a better understanding of the numerical results are becoming increasingly frequent. For instance, examples are numerous for the Delft3D model (in particular, the FLOW module), one of the most popular models for simulating hydrodynamics of natural environments. Garcia et al. (2015) and Baracchini et al. (2020b) adopted a derivative-free algorithm for non-linear least squares (DUD from Ralston and Jennrich (1978)) within OpenDA (El Serafy et al., 2007) to optimize calibration and data assimilation in their lake models; Schwindt et al. (2022) assessed the uncertainty of mixing-related model parameters through a Bayesian calibrator combined with a Gaussian process emulator; Xia et al. (2021, 2022) developed the PODS tool (Parallel Optimization with Dynamic coordinate search using Surrogates) and showed the potential of using surrogate models to accelerate global optimization. A surrogate model, also known as emulator, is a simplified and computationally efficient empirical model (Castelletti et al., 2010) that mimics the behavior of a computationally expensive model based on real model training data. In the search for the optimal parameters, the largest computation cost is indeed related to the real model runs. Thus, fast surrogate model runs are alternated with real model runs such that a lower number of real model evaluations is needed. Surrogate models can be included in traditional optimization tools such as Bayesian calibration (Ma et al., 2024) and are highly effective in speeding up the calibration process.

Many opportunities are offered by automated calibrators, saving time and finding global optimal solutions, and many other optimization algorithms have been applied for the calibration of hydrodynamic models. We refer to Xia et al. (2021) for a review

of all these aspects. However, much care must be taken when model parameters are calibrated, either manually or automatically. ~~Many examples in literature indeed report optimal values laying in~~ The literature reports examples where calibration has guided modelers in regions of the ~~parameters~~ parameter space that hold no physical ~~relevance~~meaning (e.g. Baar et al. (2019), well discussed by Tritthart et al. (2024)). Such unrealistic parameters are often justified as those giving the best model performance, but it is well known that (i) ~~a~~ different combination of parameters might give the same result and (ii) the performance of the model heavily ~~depend~~ depends on what metric is adopted. This is particularly true for hydro-thermodynamic models calibrated on a single variable, e.g. temperature profiles only (Amadori et al., 2021; Xia et al., 2022).

The objective of this work is to show how the use of automatic calibration can help identify flaws in the structure of even state-of-the-art models. Analyses of how these tools can foster a better ~~comprehension~~ understanding of the numerical results have started to appear in the literature~~, e.g., the recent example of Bayesian calibration of a Delft3D model (Schwindt et al., 2022). Here we adopt a similar approach to~~ . In the previously cited work by Schwindt et al. (2022), the authors were able to identify unrealistic model setups from the high uncertainty of the a posteriori distribution of mixing-related parameters. With a similar philosophy, here we evaluate the parameterization of the heat distribution along the water column as coded in Delft3D. In particular, we focus on the ~~Secchi depth (D_s) parameter . The importance of properly accounting for water~~depth of the Secchi disk (Secchi, 1864, hereafter denoted as D_s or simpy "Secchi depth"), as a parameter representing the transparency of the water. Considering water transparency in hydro-thermodynamic models has ~~been long recognized. Such a quantity enables to improve~~ long been recognized as fundamental to enhance the prediction of water temperature ~~, eventually accounting for biologically-related mechanismseven if those~~ (e.g. Henderson-Sellers, 1986; Zolfaghari et al., 2017; Shatwell et al., 2019), as it allows the inclusion of biologically driven mechanisms, even when they are not explicitly parameterized~~(e.g. reduction of transparency in occurrence with algal blooms, Rahaghi et al. (2024)). This aspect becomes even more relevant in data-scarcity contexts, where satellite-derived estimates of extinction coefficients can greatly improve the performance of even simplified models (e.g. Zolfaghari et al., 2017).~~

. D_s is used to model the distribution of the incoming solar ~~shortwave~~ short-wave radiation from the surface to deeper layers depending on how deep it can penetrate the water column. ~~Despite being~~Although it is often considered a calibration parameter ~~of the model~~ in models (e.g. Wahl and Peeters, 2014; Soulignac et al., 2017; Piccioni et al., 2021; Xia et al., 2021), D_s is actually a measurable quantity ~~and can significantly vary in time . Experience from previous applications of~~ that can vary significantly over time and, in some cases, in space. In large lakes, differences between onshore and offshore D_s measurements can be of the same order of magnitude as seasonal variations (e.g. Pothoven and Vanderploeg (2020) found up to 4 times lower transparency in coastal waters compared to pelagic areas of Lake Michigan). Massive algal blooms, combined with basin-scale circulation, can also drive complex spatial patterns of water transparency as those observed by Rahaghi et al. (2024) in Lake Geneva, with D_s ranging between < 2 and $6 >$ m on the same day.

However, experience from previous Delft3D ~~in different lakes shows however~~ applications in various lakes shows that the calibrated value of D_s generally resides in the lower range of measured values (Wahl and Peeters, 2014; Soulignac et al., 2018), or it is even ~~smaller than~~ lower than the observed values (~~e.g.~~ for example, Amadori et al. (2020) and Piccolroaz et al. (2019) rescaled the observed Secchi depth by a factor of 0.5). Through the ~~automatic calibrator DYNO-PODS~~ DYNO-PODS

automatic calibrator (Xia et al., 2022), we extend the search for the optimal D_s in a broader parameter space and explore uncharted regions for manual calibration. We demonstrate how this sheds light on a flaw in the heat flux parametrization of Delft3D and let us improve the physics behind such a widely used model with a small, yet significant, modification to its source code.

In the following sections, we first formulate the problem (Section 2). Here we summarize how Secchi depth is used in Delft3D (2.1), how surface warming is generally simulated in models (2.2), what are the physical implications of the current Delft3D parametrization (2.3) and of the modified version we implemented (2.4). In Section 3, we introduce the calibration strategy and numerical experiments. In Section 4, we first show the results of the original formulation of the Delft3D heat flux scheme (4.1). We then present the results of the calibration tests that point out the model's unexpected behavior (4.2). Finally, we show the gain in physical reliability achieved after our modification of the Delft3D source code (4.3).

2 Formulation of the problem

2.1 Use of Secchi depth in models

The three-dimensional model Delft3D-FLOW (Lesser et al., 2004) numerically solves the Reynolds Averaged Navier Stokes equations (RANS) under the Boussinesq and shallow-water assumptions. The horizontal velocity components and the water surface are solved by integrating the continuity and horizontal momentum equations, while the vertical velocities are computed from the continuity equation. The transport of heat is modeled by an advection-diffusion equation, assuming that there is no heat exchange at the lake bed.

In the heat equation module of the model, all the heat flux terms are estimated with empirical formulas. Here, we focus on the penetration of solar short-wave radiation, which is modeled with a reduction of the energy flux per unit area (Wm^{-2}) along the water column, as in the commonly used Beer law:

$$H_{\mathrm{sw}}(z) = H_{\mathrm{sw0}} \exp\left(-\gamma z\right), \tag{1}$$

where H_{sw0} is the downward short-wave radiation at the water surface (already considering the effect of albedo), $H_{\mathrm{sw}}(z)$ is the energy flux that reaches the depth z, and γ is the extinction coefficient (m^{-1}). This simplified model refers to the light attenuation coefficient, which can be easily measured in situ, and is appropriate for the region of the visible spectrum (380-750 nm). Accordingly, the extinction coefficient γ is defined based on the depth D_s (m) at which the Secchi disk remains visible, through the simple relation:

$$\gamma = \frac{c_\gamma}{D_s}, \tag{2}$$

where the value $c_\gamma = 1.7$ is used as standard in Delft3D (Deltares, 2023) and was originally proposed by Poole and Atkins (1929).
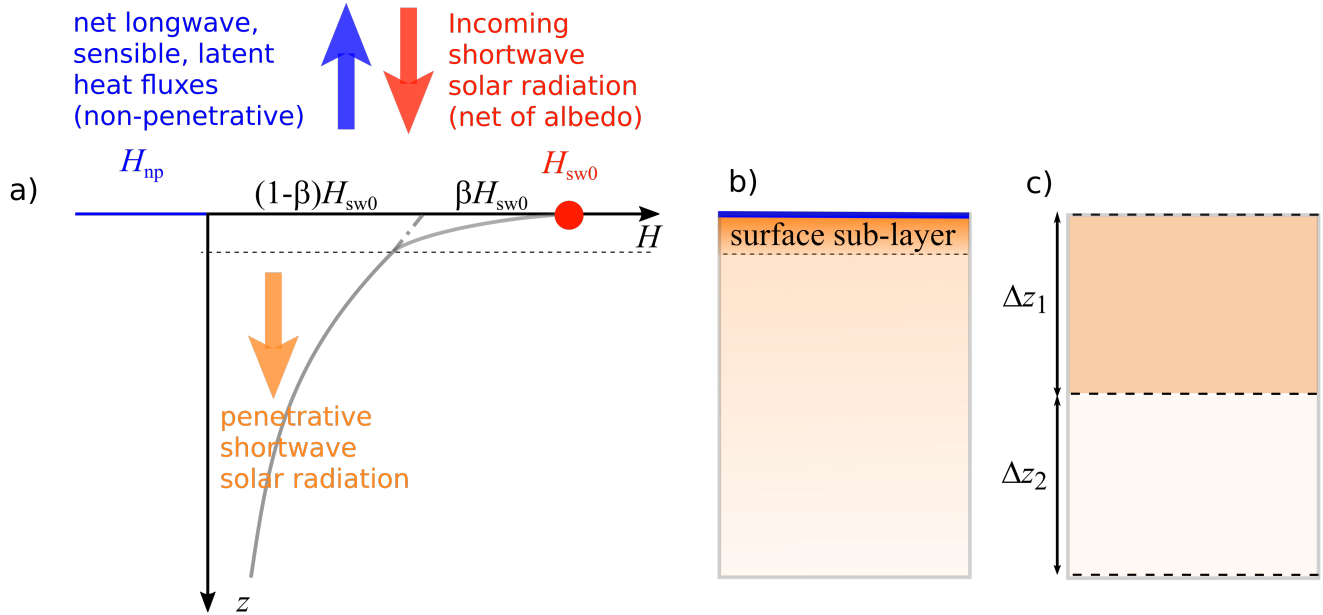
Figure 1. a) ~~Radiative~~ Heat transfer at water surface. Non-penetrative terms H_{np} are lost at the air-water interface (blue line, representative of the skin layer), while a sub-layer at the surface retains a fraction β of incoming solar radiation H_{sw0}, and only the fraction $(1-\beta)H_{sw0}$ of the solar ~~shortwave~~ short-wave radiation penetrates along the water column. The two grey lines represent the exponential decay of incoming heat flux in the surface sub-layer (above the horizontal dashed line) and beneath it. b) Conceptual illustration of how the process in a) evolves along the water column, with orange representing heat penetrating from the surface to the deeper layers; c) schematic of how a) is then parameterized in a vertically layered computational grid.

However, the actual penetration of the ~~shortwave~~ short-wave radiation depends on the different wavelengths ~~composing~~ that make up its spectrum. In principle, each wavelength has a different extinction coefficient (Zaneveld and Spinrad, 1980). Much of the ~~shortwave~~ short-wave radiation energy, in particular longer wavelengths in the ~~near to shortwave-infrared~~ near-infrared to short-wave-infrared region (750-2500 nm), is absorbed at the water surface regardless of its optical properties. The latter are indeed related to the concentration of constituents and ~~mostly~~ mainly affect the extinction coefficient in the visible region (400-700 nm).

Among many formulations available to properly describe the penetration of downward solar radiation in the water column (Morel and Antoine, 1994), the simplified version reported in Henderson-Sellers (1986) is commonly applied in many numerical models:

$$H_{sw}(z) = (1-\beta)H_{sw0}\exp(-\gamma z).\qquad(3)$$

In such a formulation, β is the fraction of the ~~shortwave radiation absorbed~~ short-wave radiation absorbed in a region close to the surface.

As a consequence, the surface layer is warmed up by a localized total heat flux,

$$H_{\text{surf}} = H_{\text{np}} + \beta H_{\text{sw0}}, \tag{4}$$

which is the sum of the absorbed short-wave radiation βH_{sw0} and of the other non-penetrative fluxes H_{np} (see Figure 1). The non-penetrative heat flux,

$$H_{\text{np}} = H_{\text{lw}}^{\downarrow} - H_{\text{lw}}^{\uparrow} \pm H_{\text{sens}} \pm H_{\text{lat}}, \tag{5}$$

is the result of the net downward long-wave radiative flux $H_{\text{lw}}^{\downarrow}$ (from the sky to the lake), the upward long-wave radiation $H_{\text{lw}}^{\uparrow}$ (emitted by the lake surface), and the sensible and latent heat fluxes, H_{sens} and H_{lat} respectively. All these terms are associated with a generic surface heat flux H_{surf}, but it is worth noting that non-penetrative terms act just at the interface between water and air, while βH_{sw0} is absorbed in a layer of the order of tens of centimeters (Henderson-Sellers, 1986) (Fig. 1b). From the viewpoint of numerical modeling (Fig. 1c), the fraction β of the incoming solar radiation represents a source of heat in a shallow layer of water, while non-penetrative terms (usually negative) represent a sink of heat, as will be shown later. All these fluxes are normally parameterized in only one computational cell, i.e. the one immediately below the water surface, where the contributions of both equations 3 and 4 should be accounted for.

Modified versions of the Beer law were introduced in some lake-dedicated models to account for the different absorption of heat depending on the spectral bands of solar radiation. This is the case of the General Lake Model (Hipsey et al., 2019), where the authors attribute 55% of the incident solar radiation to near-infrared (NIR) and ultraviolet (UVA, UVB) radiation heating the surface directly. A default value of 45% for β is implemented in CE-QUAL-W2 (Cole and Wells, 2015), but different values of 24 to 69% are recommended depending on the type of environment, with higher values attributed to pure waters (63%) and coastal waters (69%) and lower values (24 to 58%) to lake waters. A similar approach with different proportions (35% for NIR, 65% for PAR, UVA, UVB) was adopted by Thiery et al. (2014) to simulate Lake Kivu thermal structure with an ensemble of different lake models. Among these, models explicitly including β are SimStrat (Goudsmit et al., 2002), LAKEoneD (Joehnk and Umlauf, 2001), LAKE (Stepanenko and Lykossov, 2005), MINLAKE96 (Fang and Stefan, 1996).

Although accounting for this fraction of heat at the surface is important in lakes, to the best of our knowledge, the absence of this term in Delft3D never raised any concern in its previous applications to lacustrine environments. The reason is that the extinction coefficient γ estimated with equation (2), or equivalently the depth of the Secchi disk D_{s}, is considered a calibration parameter, even if D_{s} is a measurable quantity.

2.2 General formulation of surface layer warming

Hydro-thermodynamic models such as Delft3D solve a thermal energy balance, where the heat source is usually dominated by the heat exchanged at the air-water interface. The relevant terms in the equation for the transport of heat can be represented

as follows:

$$\frac{\partial T}{\partial t} + \text{advection} + \text{diffusion} = -\frac{\partial \phi}{\partial z}, \tag{6}$$

where T is water temperature and ϕ is the vertical heat flux. We assume that the vertical coordinate z is pointing downwards, such that the heat flux is positive downwards (consistent with the direction of solar radiation). The source term on the right-hand side of the equation represents the local heating, which depends on the absorption of the radiative heat flux. Without this source term, the advection and diffusion terms can only redistribute the heat in the domain, so that the total heat content of the lake is conserved if no flux is exchanged at the boundaries.

In equation (6), we have simplified the notation by introducing the flux (units $\mathrm{K\,s^{-1}\,m}$)

$$\phi = \frac{H}{\rho\,c_p}, \tag{7}$$

which scales the total heat flux H with water density ρ (units $\mathrm{kg\,m^{-3}}$) and specific heat capacity at constant pressure c_p (units $\mathrm{J\,kg^{-1}\,K^{-1}}$), assuming them to be constant, as a first approximation. Hereafter, the definition of ϕ_x based on equation (7) is adopted for all heat fluxes H_x, where the subscript x indicates the component.

The net heat flux at the air-water interface is

$$\phi_{\text{net}} = \frac{H_{\text{net}}}{\rho\,c_p} = \phi_{\text{sw0}} + \phi_{\text{np}} = (1 - \beta)\,\phi_{\text{sw0}} + \phi_{\text{surf}}. \tag{8}$$

In a depth-averaged model, the net heat flux is responsible for the change of the whole-lake temperature. Neglecting the sub-daily variability, the warming of the lake is related to the total heat exchanged during a day:

$$\int_{\text{diel}} \phi_{\text{net}}\,dt = \int_{\text{diel}} \phi_{\text{np}}\,dt + \int_{\text{sunlight}} \phi_{\text{sw0}}\,dt, \tag{9}$$

where the integrals are defined on a 24-hour period (covering the periods of sunlight, with $\phi_{\text{sw0}} > 0$, and night, with $\phi_{\text{sw0}} = 0$). The total exchanged heat, $\int_{\text{diel}} \phi_{\text{net}}\,dt$, is typically much smaller than the two individual components on the right-hand side of equation (9), as can be seen in Figure 2a.

Under these conditions, and to illustrate the behavior of the system, it can be assumed that $\int_{\text{diel}} \phi_{\text{net}}\,dt \simeq 0$. Therefore, the non-penetrative heat lost through the surface must approximately balance the solar heating, i.e. $\int_{\text{diel}} \phi_{\text{np}}\,dt \simeq -\int_{\text{sunlight}} \phi_{\text{sw0}}\,dt$, so that the non-penetrative components of the heat flux have a cooling effect in most cases. In terms of diel averaged values (indicated using angle brackets), the condition can be written as

$$\langle \phi_{\text{np}} \rangle \simeq -\frac{1}{\tau_{\text{day}}} \int_{\text{sunlight}} \phi_{\text{sw0}}\,dt < 0, \tag{10}$$
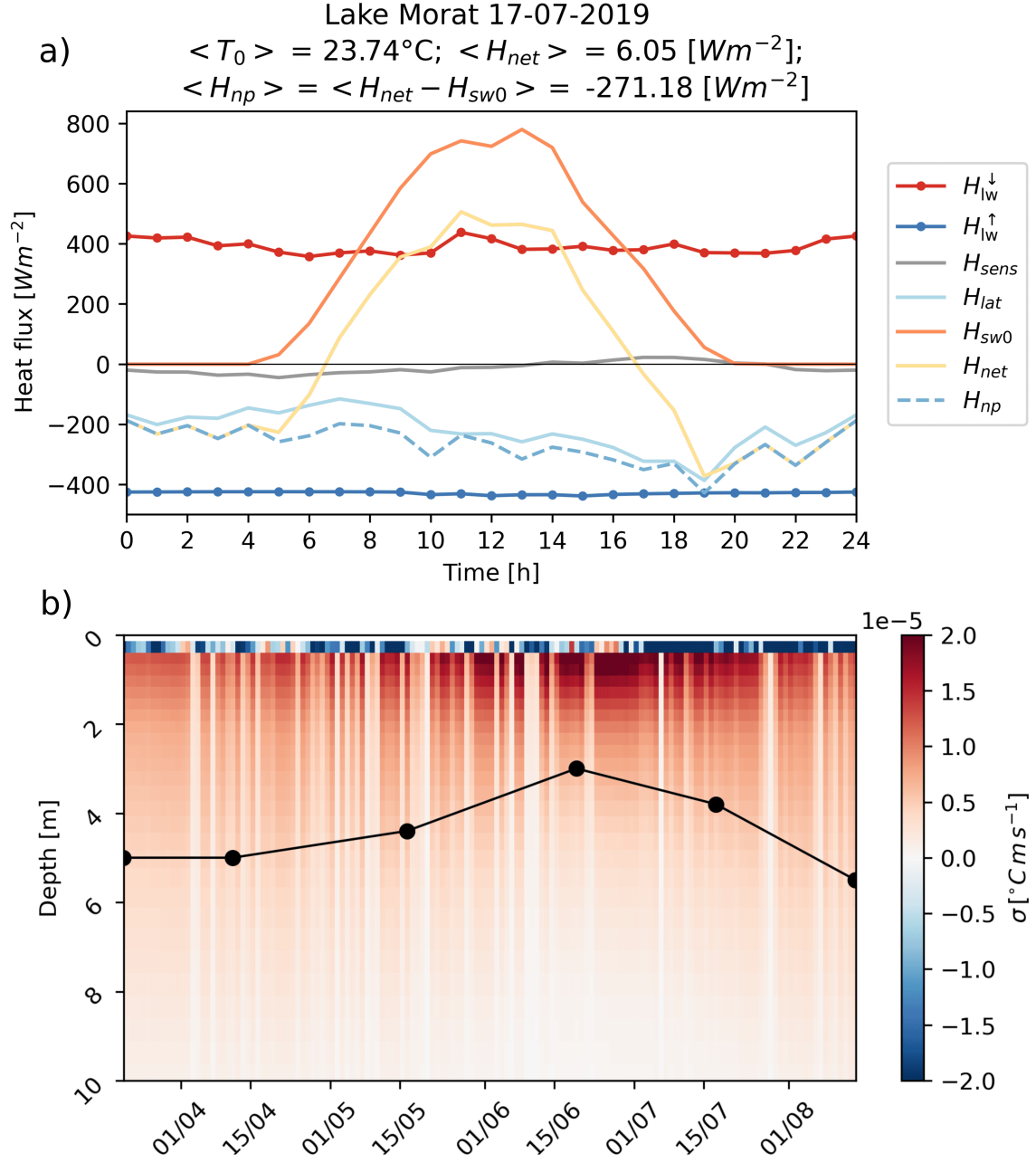
Figure 2. a) Heat fluxes computed using the meteorological forcing and the simulated surface water temperature from the base setup model ($\mathcal{O}$ in Table 1) for Lake Morat on a sample summer day, i.e. 17 July 2019. b) Warming rates simulated along the water column calculated according to equations (14)-(15) using the flux calculated by model $\mathcal{O}$. Profiles are displayed at 12:00 UTC. The black line represents the measured Secchi depth (D_s) considered in the simulation, with black dots indicating the exact measurement day. Depths are limited to 10 m to improve visibility of the surface layer.

where τ_{day} is 24 hours. Since the sub-daily variability of the non-penetrative fluxes is small compared to the solar radiation flux, it follows that $\phi_{\mathrm{np}} \simeq \langle\phi_{\mathrm{np}}\rangle$ (see, e.g., the example in Figure 2), and equation (8) implies that lakes are typically subject to (relative) cooling during night and to (relative) warming during sunlight.

195   The numerical model solves the discretized form of the heat equation (6) in a ~~layer~~ $i - th$ layer of thickness Δz_i. ~~Neglecting~~ If the advective and diffusive fluxes ~~for this simplified example, it reads~~ are neglected, the variation in time of the temperature T_i^* (where * denotes the idealized case of no advection and diffusion) can be written as:

$$\frac{\Delta T_i^*}{\Delta t} = \frac{\phi_{i-1/2} - \phi_{i+1/2}}{\Delta z_i} = \sigma_i, \qquad (11)$$

whereby we define the warming rate σ_i (units ~~). Note that T^* here refers to the theoretical temperature for a lake where~~

200 ~~advective and diffusive terms are neglected. In fact~~Ks^{-1}). In fact, assuming that the horizontal dimensions of the grid remain vertically constant and in the absence of advection and diffusion, the temperature T_i^* of the computational cell i ~~(the vertical index, assuming that the horizontal dimensions of the grid do not change vertically) increases~~ increases only if the heat flux entering the cell exceeds the flux leaving the cell. When the heat source comes from above, this means that warming occurs if the heat flux entering from the upper face ($i - 1/2$, for a staggered Arakawa C-grid) is larger than the flux leaving the cell from

205 the lower face ($i + 1/2$).

Indicating with $i = 1$ the computational cell at the top of the water column (see also Figure 1c), the flux at the upper boundary condition of the model is $\phi_{1/2} = \phi_{\mathrm{net}}$, while the flux at the bottom of the first computational cell is $\phi_{3/2} = (1 - \beta)\phi_{\mathrm{sw0}}\exp(-\gamma\Delta z_1)$. Hence, the warming rate in the uppermost computational layer ~~in equation (12)~~ can be obtained by combining equations (3),(4),(7),(8) into equation (11):

210
$$
\begin{aligned}
\sigma_1\Delta z_1 &= (\phi_{\mathrm{sw0}} + \phi_{\mathrm{np}}) - (1 - \beta)\phi_{\mathrm{sw0}}\exp(-\gamma\Delta z_1) \\
&= \beta\,\phi_{\mathrm{sw0}} + \phi_{\mathrm{np}} + (1 - \beta)\phi_{\mathrm{sw0}}\left[1 - \exp(-\gamma\Delta z_1)\right]. \qquad (12)
\end{aligned}
$$

Extending the same approach to the layers below the one at the surface, we can express the warming rate in layer $i = 2$ as

$$
\begin{aligned}
\sigma_2\Delta z_2 &= (1 - \beta)\phi_{\mathrm{sw0}}\exp(-\gamma\Delta z_1) - (1 - \beta)\phi_{\mathrm{sw0}}\exp\left[-\gamma(\Delta z_1 + \Delta z_2)\right] \\
&= (1 - \beta)\phi_{\mathrm{sw0}}\exp(-\gamma\Delta z_1)\left[1 - \exp(-\gamma\Delta z_2)\right]. \qquad (13)
\end{aligned}
$$

215 During sunlight ($\phi_{\mathrm{sw0}} > 0$), the second term on the right hand side of equation (13) (first line) is smaller than the first one, ($\Delta z_1 + \Delta z_2 > \Delta z_1$). Hence, the layer $i = 2$ is always warmed. Similar considerations apply to all the other layers with $i > 1$. ~~On the contrary~~In contrast, σ_1 may be negative (producing local cooling of the surface layer) even during warming periods, depending on the value of β and ϕ_{np}.

The difference in the warming rates of layers 1 and 2 as predicted by ~~equation~~ equations (12) and (13) may produce a

220 vertical temperature gradient, which tends to be balanced by the vertical diffusive flux in equation (6). The effectiveness of such a flux in balancing the differential warming highly depends on the intensity of the turbulence. In particular, increasing the eddy diffusivity could make the entire profile more homogeneous. As ~~we will discuss~~ discussed in the following sections, turbulence might compensate the effect of different schemes for the penetration of solar radiation, for instance if $\beta = 0$, as in the original version of the Delft3D heat equation module.

2.3 Case with $\beta = 0$

The original version of Delft3D, which adopts the Beer law as in equation (1), can be described as a particular case of the scheme described above in which $\beta = 0$. In this case the warming rate of the top layer ($i = 1$) expressed by equation ~~(14)~~ (12) can be simplified as in (14):

$$\sigma_1 \Delta z_1 = \phi_{np} + \phi_{sw0}\left[1 - \exp\left(-\gamma \Delta z_1\right)\right]. \tag{14}$$

When γ is small (~~D_s very large~~ very large D_s, i.e., very transparent water), the last term within square brackets ~~tends to vanish, so that the warming of the surface layer is $\sigma_1 \simeq \phi_{np}/\Delta z_1$, which is typically negative because $\phi_{np} \simeq \langle\phi_{np}\rangle < 0$, as~~ becomes less important and non-penetrative terms prevail. As discussed in section 2.2, $\phi_{np} \simeq \langle\phi_{np}\rangle < 0$ (see also Figure 2a). Therefore, the layer $i = 1$ may cool down also when sunlight is present ($\phi_{sw0} > 0$)~~, while the lower layers are always warmed up by the penetrative shortwave radiation ,~~. The lower layers instead always receive heat, despite being maybe small in very transparent conditions, as the warming rate σ_2 only depends on penetrative short-wave radiation as implied by equation ~~(15,~~ (15), which is obtained by setting $\beta = 0$ in equation 13~~)~~.

$$\sigma_2 \Delta z_2 = \phi_{sw0} \exp\left(-\gamma \Delta z_1\right)\left[1 - \exp\left(-\gamma \Delta z_2\right)\right]. \tag{15}$$

2.4 Case with $\beta > 0$

We modified the Delft3D source code by substituting the Beer law (1) with the complete version (3), as reported in Appendix A. Hence, the scheme reported in equations (12) and (13) with $\beta > 0$ applies. Focusing on the warming of the top layer and repeating the same argument as for equation (14), in case of small γ the last term of equation (12) vanishes and $\sigma_1 \simeq (\beta\,\phi_{sw0} + \phi_{np})/\Delta z_1$. Now, the sum of the two terms is not necessarily negative, as $\beta\,\phi_{sw0}$ can balance the negative contribution of ϕ_{np}, and the surface layer is not forced to cool down (or not so much) during sunlight. A proper choice of the coefficient β ~~guarantees a physically sound calibration of the numerical~~ can improve the physical consistency of the calibrated model.

3 Material and methods

Lake Morat, Switzerland~~is chosen as a~~, is the test site for our study. This is one of the alpine lakes included in the AlpLakes network[1] for which a calibrated setup of the Delft3D model is available. The details of the case study and the base Delft3D setup have been explained in Appendix B.

3.1 Calibration strategy

For the calibration of the Delft3D model we adapted to our needs the procedure DYNO-PODS based on a parallel surrogate global optimization method (Xia et al., 2021, 2022), which allows for calibrating multiple parameters simultaneously. DYNO-PODS runs parallel simulations and finds the best result at each iteration based on a cost function ~~ϵ_k to be minimised~~ε_k to be

---

[1]https://www.alplakes.eawag.ch

minimized. We refer to the DYNO-PODS documentation and in particular to Xia et al. (2022) for an exhaustive description of the surrogate model implemented in DYNO-PODS.

255   The purpose of optimization is to ~~correclty~~ correctly reproduce measured water temperature in one single station. We therefore tested two objective functions: (i) an error function ~~ϵ_T~~ ε_T based on the temperatures measured ~~along the entire~~ throughout the profile, as in equation (16); (ii) an error function ~~ϵ_{T0}~~ ε_{T0} based on the temperatures measured only at the lake surface~~only~~, as in equation (17).

    The distributed error ~~ϵ_T~~ ε_T is estimated by the average root mean square deviation

$$260 \quad \underline{\epsilon}\varepsilon_T = \sqrt{\frac{1}{N_t}\sum_{j=1}^{N_t}\frac{1}{N_z}\sum_{i=1}^{N_z}\left(T_{i,j}-\widehat{T}_{i,j}\right)^2}, \tag{16}$$

where N_t is the number of temporal profiles, N_z is the number of measuring points along the vertical, $T_{i,j}$ is the observed temperature at the depth z_i and time t_j, and $\widehat{T}_{i,j}$ is the corresponding value simulated by the model. If ~~ϵ_T~~ ε_T is chosen as objective function, the model is ~~optimised~~ optimized towards the best representation of both surface and deep water temperatures.

    The error at the surface ~~ϵ_{T0}~~ ε_{T0} is estimated via the root mean square deviation of the surface temperature T_0:

$$265 \quad \underline{\epsilon}\varepsilon_{T0} = \sqrt{\frac{1}{N_t}\sum_{j=1}^{N_t}\left(T_{0,j}-\widehat{T}_{0,j}\right)^2}. \tag{17}$$

If ~~ϵ_{T0}~~ ε_{T0} is chosen instead as objective function, the calibrated model provides the best representation of surface temperature, regardless of the performance on water temperature below.

3.2   Calibration tests

Starting from a base setup ($\mathcal{O}$ in Table 1, for more details~~see~~, see Appendix B), we tested how the performance of the model
270 could be improved by calibrating the Secchi depth. To preserve the time-dependence of measured values of Secchi depth we introduced a parameter δ defined as follows:

$$\widehat{D}_{\mathrm{s}} = \delta\, D_{\mathrm{s}}\,\underline{,} \tag{18}$$

where $\widehat{D}_{\mathrm{s}}$ is the Secchi depth value used in the model and D_{s} refers to the original value measured on-site.

    We first calibrated δ with the original model ($\mathcal{O}_{cal}$ in Table 1). Two separate calibration tests were performed considering
275 the two objective functions defined in equations (16)-(17)~~(only the first value is reported in the Table 1, for more results see~~. Since the error along the entire water column (ε_T) is a standard objective function, only the details of these results are reported in Table 1. However, we also show and discuss the relevant results of tests with ε_{T0} below in Table 2 in Section 4.2).

    As a second step, we tested the effect of introducing the coefficient β in a modified version of the model, $\mathcal{M}_{cal}$, where all other parameters were maintained as in $\mathcal{O}$, and δ and β were calibrated.

280   Finally, we run a new base calibration for the modified model $\mathcal{M}$, where the full set of parameters was optimized, including β.

Table 1. Calibration tests setups with values of the main parameters (wind drag corrective coefficient α, free convection coefficient (CFrcon), horizontal eddy diffusivity and viscosity coefficients (Dicouv, Vicouv) and Ozmidov length scale (Xlo), scaling of Secchi depth (δ) and fraction of short-wave solar radiation absorbed in the first layer (β). $\mathcal{O}$ identifies the original Delft3D model, while $\mathcal{M}$ the modified version including β. Numbers are reported only when calibrated in the test, and "n.c." stands for "not calibrated". For $\mathcal{M}_{cal}$ test, not calibrated parameters are the same as in $\mathcal{O}$. "n.a." stands for "not applicable" and only refers to the β parameter which does not exist in $\mathcal{O}$ model version.

Model	Test	α	Cfrcon	Dicouv	Vicouv	Xlo	δ	β	ε_T
		-	-	$\mathrm{m^2 s^{-1}}$	$\mathrm{m^2 s^{-1}}$	mm	-	-	$^\circ C$
Original	$\mathcal{O}$	0.94	0.019	2.36	1.65	0.338	n.c.	n.a.	0.57
Original	$\mathcal{O}_{cal}$	n.c.	n.c.	n.c.	n.c	n.c.	0.84	n.a.	0.49
Modified	$\mathcal{M}_{cal}$	n.c	n.c	n.c	n.c	n.c.	1.00	0.32	0.47
Modified	$\mathcal{M}$	0.98	0.004	1.84	1.28	0.118	n.c.	0.33	0.45

All calibration tests were performed in those months when water transparency prominently influences the formation of a stratified thermal structure in dimictic lakes, i.e. the spring and summer months after winter mixing. The simulated period was 20 March to 14 August 2019. Sensitivity tests on the model response to variations in Secchi depth during the cooling season (from August onward) confirmed the assumption that it is possible to effectively determine the optimal value of δ limiting the calibration to the period when stratification occurs (not shown).

In order to gather a satisfactory convergence of the automated calibration tests, we run 10 iterations of 12 parallel simulations, for a maximum value of 120 evaluations for each calibration test. This aligns with Xia et al. (2021), who found that 120 model evaluations were necessary to achieve a solution similar (in performance) to manual calibration in their case study. The simulation time for 5 months was about 11 hours on a High Performance Computing Cluster. Each simulation utilized 6 cores on a single node, resulting in 72 CPU cores being used per iteration.

4 Results

4.1 Base setup model

In this section, we present the results of the calibration of the original model $\mathcal{O}$ (without β, see Table 1) and without any modification of the Secchi depth. This setup implicitly corresponds to $\beta = 0$ and $\delta = 1$. Figure 2a shows the heat balance terms computed on an hourly time scale at the lake-atmosphere interface on a summer day (17 July 2019), for which in situ measurements of water temperature are available. As expected from the theoretical considerations in Section 2.2, the net heat flux is heavily affected by the short-wave solar radiation H_{sw0}, which however vanishes at night, producing a strong sub-daily variability that is extremely large compared to the other terms. Moreover, if the short-wave radiation

is excluded from the overall balance H_{net}, the resulting net heat flux accounting only for the ~~non penetrative~~ non-penetrative terms (H_{np}, dashed line) is ~~negative also~~ also negative during daylight hours.

~~a) Heat fluxes computed using the meteorological forcing and the simulated surface water temperature from the base setup~~
305 ~~model ($\mathcal{O}$ in Table 1) for Lake Morat on a sample summer day, i.e. 17 July 2019. b) Warming rates simulated along the water~~ ~~column calculated according to equations (14)-(15) using the flux calculated by model $\mathcal{O}$. Profiles are displayed at 12:00 UTC.~~ ~~The black line represents the measured Secchi depth (D_s) considered in the simulation, with black dots indicating the exact~~ ~~measurement day. Depths are limited to 10 to improve visibility of the surface layer.~~

By using the heat fluxes internally computed in the calibrated Delft3D model, we estimated the local warming rate of the
310 surface layer according to equation (14), i.e. neglecting advection and diffusion. In this simplified scheme, the heat flux at the surface resulted ~~to be~~ in an almost always negative value indicating a predominantly heat loss term in the heat equation of the surface layer. As shown in Figure 2b, where the warming rate σ is plotted everyday at noon, this happens even in the warming periods (spring-summer), with a few exceptions during extremely calm (i.e. low wind) and sunny days. ~~At~~ In layers beneath the surface, σ (estimated locally with equation (15)) ~~has instead~~ instead has a positive sign. ~~Shortwave radiation indeed~~ In fact,
315 short-wave radiation penetrates the water column with a decreasing warming rate as depth approaches the Secchi depth D_s. Hence, excluding advective and diffusive terms from the simulation as in equation (11), would result in continuous cooling of the surface layer and slight warming of deeper layers. Although this is not the case in the Delft3D model, which accounts for advective and diffusive fluxes, the tendency can ~~be noted also~~ also be noted in the numerical results, as ~~it~~ will be shown below (Figure 4).

320 ## 4.2 Optimal scaling of Secchi depth

The results of the calibration tests $\mathcal{O}_{cal}$ and $\mathcal{M}_{cal}$ ~~for assessing~~ to assess the optimal scaling ~~for~~ of the Secchi depth are reported in Table 2. The search for the optimal value of δ depends on the objective function used in the calibration tests.

In the original model $\mathcal{O}_{cal}$ (Figure 3 a-b), if the ~~target of the calibration~~ calibration target is the surface water temperature ($~~\epsilon_{T_0}~~\varepsilon_{T_0}$ in equation (17)), the optimal scaling for the Secchi depth unrealistically converges to $\delta = 0.12$. Smaller values of
325 δ lead to warmer surface temperature because γ becomes very large (low transparency; see equation (2)) and a large portion of the ~~shortwave~~ short-wave solar radiation is absorbed in the top surface layer. If the target is ~~water temperature over the~~ ~~entire water column (ϵ_{T_z}~~ the temperature of the water throughout the water column (ε_T in equation (16)), the optimal value is $\delta = 0.84$.

The same plots from ~~$\mathcal{M}_{cal}$ tests~~ the tests $\mathcal{M}_{cal}$ (Figure 3 d-f), suggest that the water temperature can be simulated without the
330 need ~~of drastically minimizing~~ to drastically minimize δ. The optimal δ converges to 0.57 and 1 for the two objective functions $~~\epsilon_{T_0}~~\varepsilon_{T_0}$ ~~and ϵ_{T_z}~~ and ε_T, respectively. The first value (0.57) will be commented on in the discussion section. Interestingly, Figure 3e-f show that the value of β converges in a physically meaningful range and the optimal value is 0.32 for both objective functions (Table 2), fully consistent with ~~literature values~~ the range of literature values (20-60%) presented in Section 2.1.
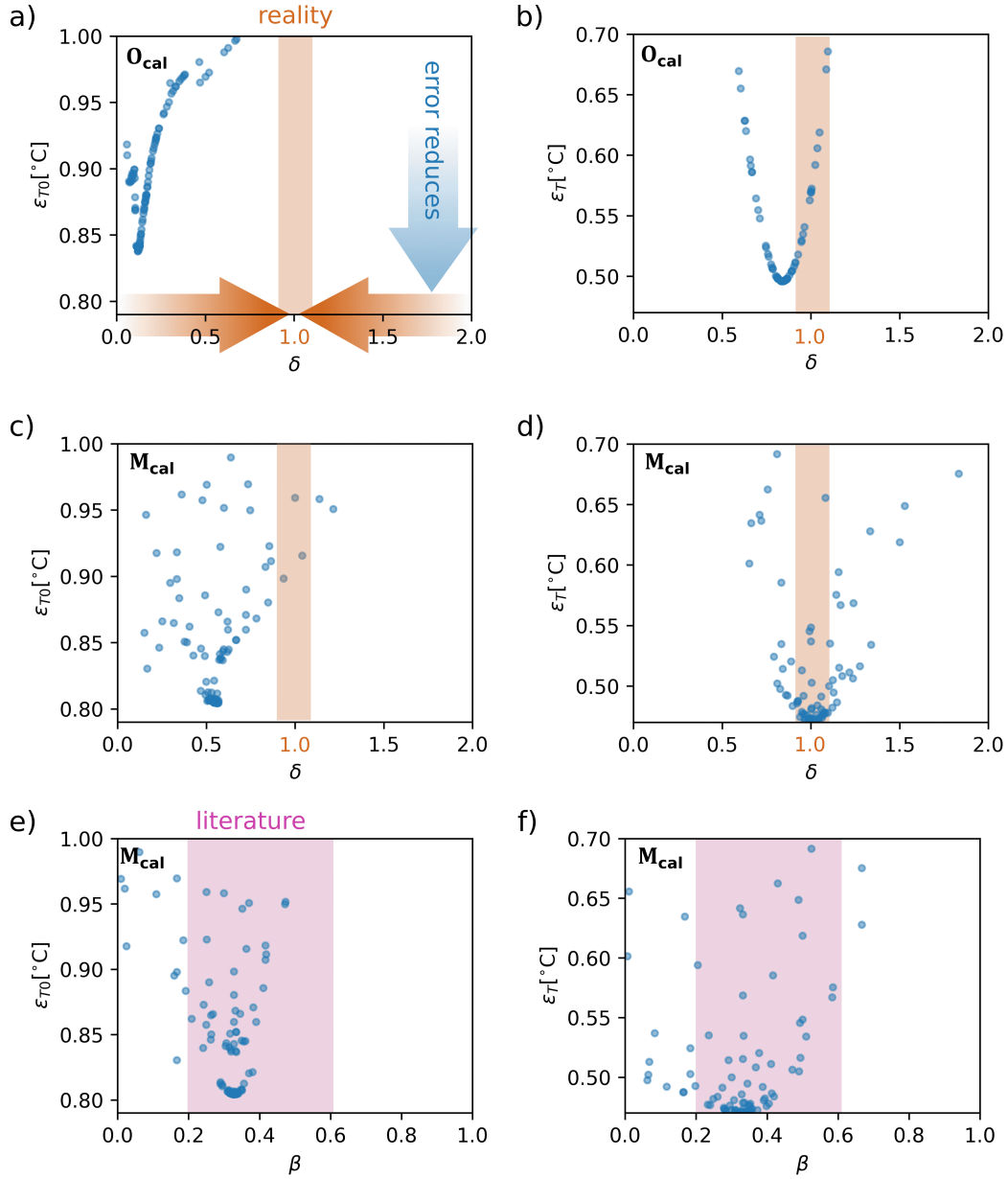
Figure 3. Convergence of different calibration tests and objective functions (a, c, e: ϵ_{T0}, error on surface temperature; b, d, f: ϵ_{Tz}, error on temperature along the water column). a-b): Calibration test $\mathcal{O}_{cal}$ for optimal δ in the original model $\mathcal{O}$ (β implicitly equals 0); c-f) calibration of both δ (c,d) and β (e,f) in $\mathcal{M}_{cal}$ tests with the modified model. The blue and orange arrows in panel a) show the main expected directions of the optimization problem in the calibration of δ (a-d). The orange shaded rectangle in panels a-d highlights the range of realistic values of δ, i.e. around 1. The pink shaded rectangle in panels e-f highlights the range of values reported for β in the literature, i.e. 0.2-0.6.

Table 2. Calibrated parameters δ and β and corresponding objective functions. "n.a." stands for "not applicable".

Test	parameter		objective function	
	name	value	name	value
$\mathcal{O}_{cal}$	δ	0.12	ε_{T0}	0.83 °C
	β	n.a.		
	δ	0.84	ε_T	0.49 °C
	β	n.a.		
$\mathcal{M}_{cal}$	δ	0.57	ε_{T0}	0.8 °C
	β	0.32		
	δ	1.00	ε_T	0.47 °C
	β	0.32		

4.3 Effect of β

335 To understand what actually changes when ~~introducing~~ the parameter β ~~in the heat flux parameterization~~is introduced in the parameterization of the heat flux, Figure 4 shows the comparison between the fully calibrated models $\mathcal{O}$ and $\mathcal{M}$ (Table 1). Panels a,b show the simulated temperature profiles on a sample summer day in the two optimal simulations ~~minimizing ε_{T_z}~~that minimize ε_T, with and without β. The two thermal profiles appear ~~very similar and consistent with~~ to be very similar to each other and are both consistent with the measurements. Both simulations correctly capture the thermocline depth and the

340 temperature of the surface well-mixed layer (about 7 m thick). However, in the model $\mathcal{O}$ (panel a) the temperature of the surface layer (23.27 °C) is slightly colder than that ~~at~~ of the layers below (23.35°C), while the surface in-situ temperature (24.15°C) is the warmest. ~~Conversely~~In contrast, the model with calibrated β, i.e. $\mathcal{M}$ (panel b), returns 24.50° C ~~at~~ on the surface.

Although the difference in temperature is relatively small between the two ~~model versions~~ versions of the models, the physics behind such thermal profiles is actually different. In Figure 4c-e we compare the eddy diffusivity D_z as simulated by the k-~~ϵ~~ε

345 model in $\mathcal{O}$ and $\mathcal{M}$ ~~in~~ on three selected days. In the original model version, stronger turbulent vertical diffusion is simulated in the epilimnion. The largest difference is visible during daylight hours, when the ~~$\mathcal{O}$ model~~ model $\mathcal{O}$ (panel c) simulates homogeneous diffusivity of the order of 10^{-4} m^{2}s^{-1} in the upper 2 m, while surface mixing is strongly inhibited in the ~~$\mathcal{M}$ model~~ model $\mathcal{M}$ (panel d), with D_z generally lower than 10^{-6} m^{2}s^{-1}. Thus, the enhanced mixing in the epilimnion simulated by the model $\mathcal{O}$ compensates for the unrealistic distribution of the warming rate caused by the absence of β, which would

350 produce cooling of the surface layer (as Figure 2a shows). This effect is visible ~~in the whole~~ throughout the simulated period (panel e), with an overestimation of D_z of the order of 10^{-4} m^{2}s^{-1} during daylight hours and higher differences during the stratified period.
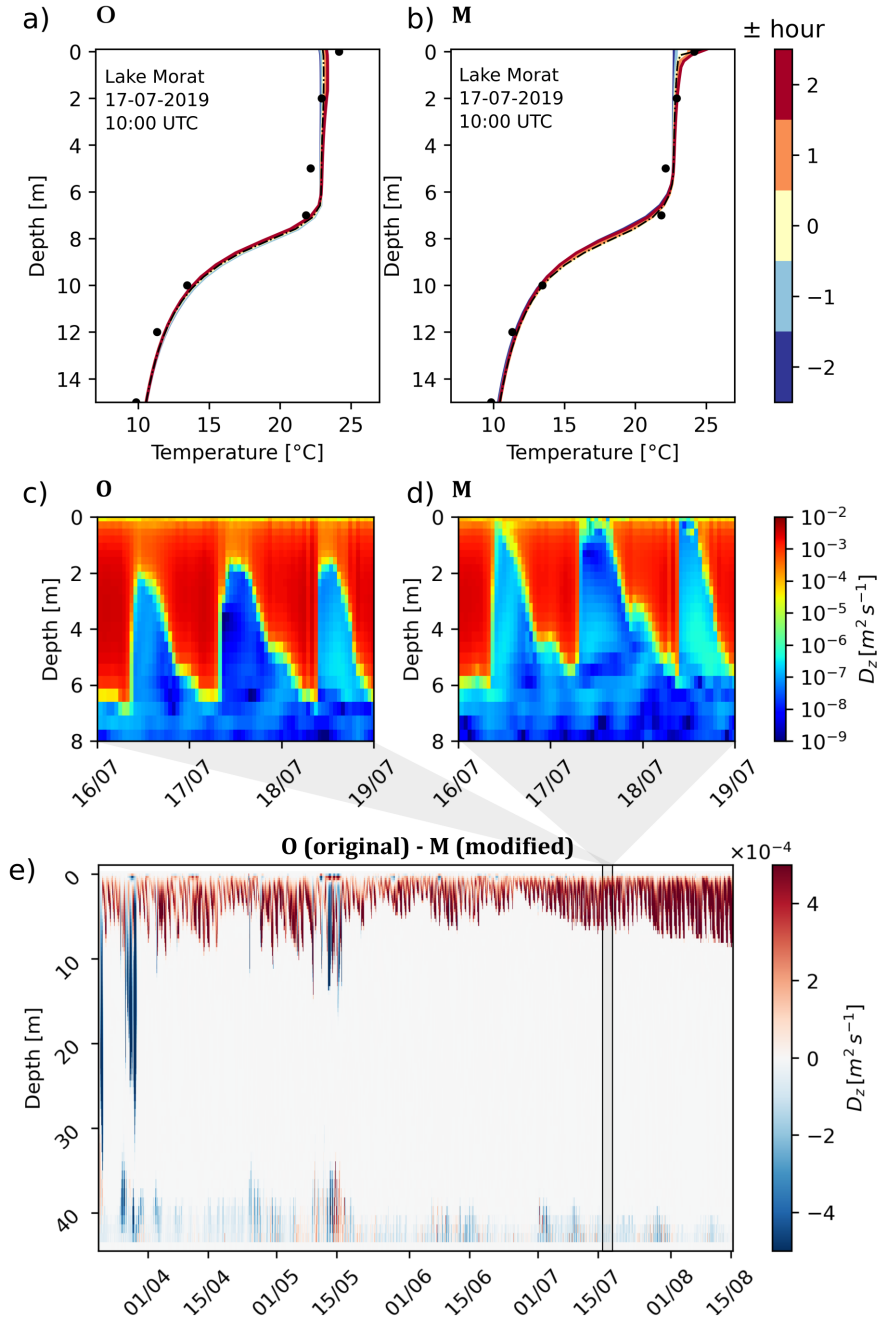
Figure 4. a-b) Temperature profiles simulated (lines) and measured (black dots) on a summer day (17 July 2019) in Lake Morat implementing the $\mathcal{O}$ and $\mathcal{M}$ model configurations with target $\epsilon_{T_z}\widetilde{\epsilon_T}$. Model results are displayed at the measurement time (10:00 am, black dashed line) and in a time range of ± 2 hours around 10:00 am. c-d) Vertical eddy diffusivity simulated in the two model setups as in a-b) ± 2 days around 17 July 2019 (black straight line in panel e). d) Difference in simulated eddy diffusivity between the two models shown on an hourly basis.

5 Discussion

The comparison between the model outputs obtained with the original ($\mathcal{O}$) and modified ($\mathcal{M}$) versions of Delft3D clearly highlights that accounting for ~~shortwave~~ short-wave absorption in a shallow layer close to the water surface improves the performance of the model and, more importantly, provides a more realistic description of the physical processes driving surface warming in lakes. The best value of the scaling parameter δ obtained with the modified model is 1 (see Table 1), which implies that the measured value of the Secchi depth in Lake Morat is a reliable input parameter for the model and does not necessarily require calibration. The fact that the measured Secchi depth can be assumed as a physically ~~meaningful~~ measurable quantity is a significant advantage in the model setup ~~with respect~~ compared to the common expectation of several Delft3D users that the Secchi depth must be calibrated (Wahl and Peeters, 2014; Soulignac et al., 2017; Piccioni et al., 2021; Xia et al., 2021). ~~Moreover,~~

Systematic calibration of model parameters, including Secchi depth, is essential not only for more accurately representing the physics of the ~~value of β estimated through the automatic calibration accurately falls within the range $\beta \in (0.3, 0.4)$ observed in different lake environments (e.g., 0.4 in Dake and Harleman, 1969) and suggested in previous modeling applications (Thiery et al., 2014; C~~ ~~.~~ system, but also for addressing the imperfections inherent in the model (e.g., neglected baroclinic mixing, imperfect turbulent schemes, unaccounted effects of tributary intrusions, inaccurate atmospheric forcing). Among all the calibration parameters, the Secchi depth is perhaps the most frequently monitored and is one of the easiest to measure, both in situ and through remote sensing. However, several previous modeling studies have highlighted significant deviations of this parameter in Delft3D from the measured values. Here, we demonstrate that this discrepancy is due to a misimplementation of certain physical phenomena in the Delft3D model, which is supported by the deviation of calibrated Secchi depths after correcting for it. However, there may still be a need to calibrate Secchi depth to account for potential errors in measurements or the sparse resolution of Secchi depth data in both time and space. Meanwhile, calibration of other parameters (e.g., wind drag, background eddy coefficients) will remain necessary to compensate for unresolved processes in the model. We believe that it is good modeling practice to avoid compensating for inaccuracies in one process by adjusting the parameters of others. We therefore suggest focusing on improving model deficiencies and promoting the correct use of parameters to ensure more reliable and physically sound results.

To what extent ~~such a~~ the modification of the heat fluxes ~~parameterization is~~ scheme was necessary to improve the ~~model performances~~ performance of the model can be discussed ~~building on the model's errors~~ based on the objective functions reported in Tables 1 and 2. Our results indicate that the base calibration of the original model $\mathcal{O}$ (Table 1) ~~can already give~~ already gives satisfactory results, with an error ~~$\epsilon_T = 0.57°$~~ $\varepsilon_T = 0.57°$C along the water column. ~~Such a~~ This result is possible ~~thanks~~ due to the automated calibration, which allows us to explore a wider range of ~~testing parameters combination compared to~~ parameter combinations compared to the manual procedure. The relatively low error also confirms the general reliability of the model with its original formulation, as demonstrated by the numerous successful lake applications in different lakes (e.g. Amadori et al., 2021; Baracchini et al., 2020a; Soulignac et al., 2017, 2018; Wahl and Peeters, 2014; Chanudet et al., 2012; Dissanayake et al., 2019, among many others). By rescaling D_s ~~through the calibration of δ ((~~ (i.e. by calibrating δ in the $\mathcal{O}_{cal}$ test),

the error reduced to 0.49°C. This result is in line with the observations of Amadori et al. (2020); Piccolroaz et al. (2019), who reduced the measured Secchi depth to improve the performance of the original Delft3D model in Lake Garda. An improvement of 0.08°C in the model error along the water column is considerable (14%), but it is achieved at the expense of the physical meaning of the Secchi depth information. If β is introduced and calibrated ($\mathcal{M}_{cal}$ test), we see that the final error $\varepsilon_{\mathcal{T}} = 0.47$°C is only slightly smaller than in the previous case, but a significant gain is obtained in the physical consistency of all inputs ($\delta = 1.0$, $\beta = 0.32$). The value of β estimated by automatic calibration falls conveniently within the range $\beta \in (0.2, 0.6)$ observed in different lake environments (e.g., 0.4 in Dake and Harleman, 1969) and suggested in previous modeling applications (Thiery et al., 2014; Cole and Wells, 2015; Schmid and Köster, 2016).

The interesting element of our analysis is that the parameter calibration allows the original model to get close to reality despite the faulty parametrization of the heat absorption at the surface. However, the price of such an optimization is the alteration of other physical processes that compensate for the flaw in the model formulation. This is evident from the enhanced diffusivity at the surface in the original model formulation (Figure 4), which derived from an unrealistic instability between the upper layer and those below due to surface cooling. This overestimation of surface mixing was also observed by Biemond et al. (2021), who compared the results of the Delft3D model in Lake Garda with in situ turbulent kinetic energy dissipation profiles. In particular, such overestimation was stronger in stratified conditions and was present also when temperature profiles were appropriately simulated.

We have shown in Figure 4 that the largest difference between the two model configurations is at the surface, and in particular during daylight hours in the stratified period. In the temperature profiles shown, a strong gradient is present between the near-surface temperature (i.e. 0 m) and the layer immediately below (2 m). On other days, when the measured surface temperature is relatively well mixed with the layers below, the difference between the two model versions is not that relevant. In fact, if enough mixing is physically provided by, for example, strong wind-induced turbulence, the model does not need to generate artificial convection to compensate for the negative buoyancy at the surface caused by the incorrect parametrization.

The thickness of the model layers also plays a role in the optimal scaling of δ. When the surface water temperature is the calibration target ($\varepsilon_{\mathcal{T}_0}$ as an objective function), the optimal value of δ is 0.57 (Figure 3c) even if β is included in the parametrization and is calibrated. Although 0.57 is definitely more acceptable than 0.12, which was obtained in the case without β (3a), it is still significantly smaller than the desirable value, i.e. $\delta = 1$. We speculate that this behavior can be interpreted by recalling that the layer in which the β fraction is absorbed is generally considered to be ~ 60 cm thick (Henderson-Sellers, 1986). This thickness is used by several models as the reference depth where the exponential decay starts (Zaneveld and Spinrad, 1980; Piccolroaz et al., 2024). Since the thickness of the surface layer in our model is 0.26 m (Table B1), it is possible that the value of δ lower than 1 again compensates for trapping the fraction β in a

~~optimizing the model to simulate surface water temperature only~~ layer too thin. It is therefore likely that if we used a thicker surface layer (e.g. 60 cm thick), the value of the parameter would be closer to 1.

Using automated tools to identify existing flaws could also help improve the representation of other simulated processes. For instance, the impact of density stratification on vertical turbulent fluxes is poorly represented in turbulent schemes, and internal wave breaking is accounted for in only a few lake models, for example Simstrat by Goudsmit et al. (2002). Models like k-ε could benefit from supervised automated calibration of parameters that are traditionally fixed based on "standard" experiments and generally poorly understood. Similarly, refining horizontal turbulent models through tuning could provide insights into whether the adopted scheme (e.g. constant coefficients or sub-grid scaling following Smagorinsky, 1963) and grid resolution are sufficient to capture the timing and location of horizontal and vertical transport processes, such as up/downwellings and gyres.

6 Conclusions

Automatic calibration tools are powerful for optimizing parameter selection in complex models, helping modelers ~~in achieving~~ achieve realistic simulations of the ~~studied environment~~ environment studied. Here, we demonstrated that such tools can also be utilized by developers to assess whether their implementations of models align with the physics, and by users to identify potential improvements needed in the modeling framework. ~~While issues with lake modelling often stem from inappropriate meteorological conditions,~~ Unlike manual calibration, which is often user-sensitive and prone to bias, automatic calibration is objective and allows one to more effectively highlight inherent flaws in model formulations. Although lake modeling issues often arise from inadequate meteorological input, automatic calibration can help users and developers ~~can leverage these toolboxes to~~ detect code-related issues~~.~~, offering a more unbiased and systematic approach. We tested the response of a hydrodynamic model, Delft3D, which is widely used by the limnological community, to the task of simulating the vertical distribution of heat in lakes. The process is regulated by the well-known Beer law describing the absorption rate of ~~shortwave~~ short-wave radiation along the water column, but its incomplete implementation in the numerical model requires a compensation mechanism ~~(via a non-physical~~ to optimally simulate surface water temperature. Such a mechanism is the nonphysical adjustment of the Secchi depth ~~instead of the including the absoprtion~~ to compensate for the exclusion of the absorption of a fraction of solar ~~shortwave~~ short-wave radiation at the surface~~) to optimally simulate surface water temperature~~. By exploiting a recently developed calibration tool (DYNO-PODS), based on surrogate models, we were able to unveil such a limitation of the Delft3D ~~parameterization~~ parametrization that was never explicitly discussed ~~,~~ and to fix it with a small modification of the source code.

In conclusion, we recommend caution ~~in~~ when blindly applying automatic tools, emphasizing the importance of evaluating the physical significance of the ~~obtained calibration parameters . We believe that the extensive use of such calibrators, and the insightful analysis of the results of the optimization , may offer similar outcomes in other applications, even in the case of computationally heavy~~ calibration parameters obtained. This calibration approach combined with detailed analysis of optimization results should yield comparable benefits in various applications, including computationally intensive simulations.

Table B1. Base setup for the Delft3D model of Lake Morat.

Nmax	Mmax	Kmax	$\Delta x, y$	Δz	Δt
-	-	-	m	m	s
41	88	77	73-110	0.26-1	60

Appendix A: Source code

455 We report here the parts of the Delft3D code which were modified to include the effect of β according to equation (3). Such modification is made in the subroutine `heatu.f90`.

At surface layer, the original source code lists:

```
qink   = corr * qsn * (1.0_fp – exp(extinc*zdown)) / extinc
```

We modified as:

460
```
qink   = corr * (1.0_fp-beta_sw) *qsn * (1.0_fp – exp(extinc*zdown)) / extinc
       + qsn * beta_sw
```

At deeper layers, the original source code lists:

```
qink   = corr * qsn * (exp(extinc*ztop) – exp(extinc*zdown))
```

We modified as:

465
```
qink   = corr * (1.0_fp-beta_sw) * qsn * (exp(extinc*ztop) – exp(extinc*zdown))
```

Appendix B: Test site and base model calibration

Lake Morat (Lake Murten in German) (Fig. B1a,b) is a Swiss lake situated at 429 m above sea level with a surface area of 22.8 km^2, an average depth of 24 m ,and a maximum depth of 45 m. In-situ water temperature profiles are sampled at monthly basis monthly in the middle of the lake (Fig. B1 b, c) over irregular depth intervals: 0 (just below the surface),2,together with
470 5,7,10,12,15,17,20,25,30,35,40 and 43. Monthly measurements of water transparency (expressed as Secchi depth, Fig. B1d) are also recorded at the same observation point.

The details on the model grid size and size of the model grid and the time spacing of the Delft3D model for this lake are reported in Table B1. As atmospheric forcing at the surface boundary of the lake we used COSMO-1 space and time varying meteorological variables (air temperature, relative humidity, wind speed, cloud cover, shortwave short-wave radiation and
475 air pressure) provided by MeteoSwiss at hourly time resolution and 1.1 km spatial resolution. The assimilated outputs, i.e., reanalysis data, based on observational measurements (Voudouri et al., 2017) were used for this purpose.

The base setup for this lake model was obtained from a year-long automated calibration implementing DYNO-PODS (Xia et al., 2022). The full set of parameters are reported in Table 1 in the main text and include the background horizontal viscosity and diffusivity coefficients (indicated as Vicouv, Dicouv in Delft3D, see Deltares 2023), the Ozmidov lenght scale (Xlo),
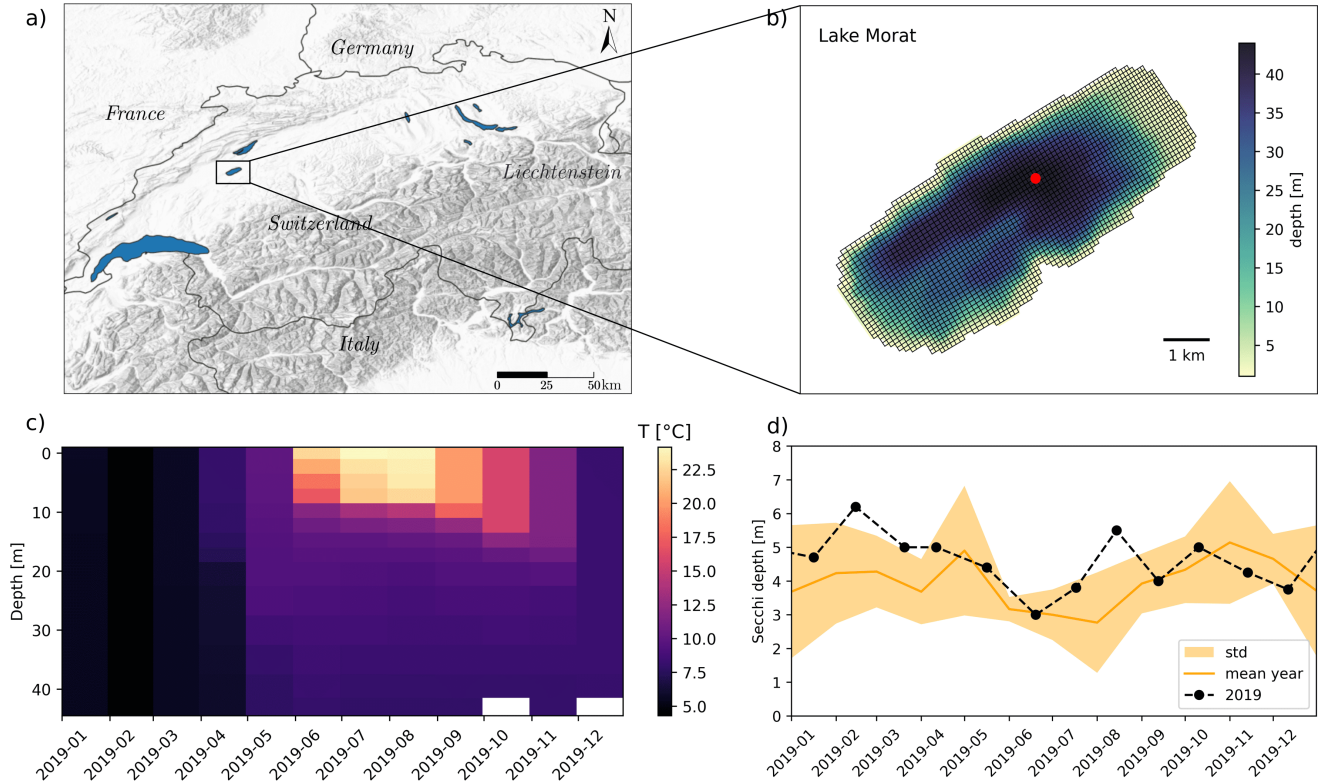
Figure B1. a) Location of Lake Morat in the alpine area; b) computational grid, bathymetry and location of in-situ monitoring station (red dot); c) heatmap of the monthly in-situ water temperature profiles for the year 2019 measured in the monitoring point; d) Time series of in-situ water transparency for 2019 (black dots) and mean year computed from the 2016-2021 time series (mean: orange thick line; standard deviation: orange shaded area).

480   the free convection coefficient (Cfrcon), and a corrective coefficient on wind drag coefficient α. For the wind drag coefficient, as in recent applications of the Delft3D model to peri-alpine lakes (e.g. Amadori et al., 2021), we took as a reference the ~~parameterization~~ parametrization proposed by Wüest and Lorke (2003) to set the coefficients of the piecewise function implemented in Delft3D: $C_{dA}^0 = 0.0044$ for wind speed lower than ~~$U_A = 0.5\,\mathrm{m\,s^{-1}}$~~ $U_A = 0.5$ ms^{-1}, $C_{dC}^0 = 0.002$ for values higher than ~~$U_C = 10\,\mathrm{m\,s^{-1}}$~~ $U_C = 10$ ms^{-1}, and linear interpolations between $C_{dB}^0 = 0.001$ and C_{dA}^0 and C_{dC}^0, respectively

485   for wind speed between ~~$U_B = 4.5\,\mathrm{m\,s^{-1}}$~~ $U_B = 4.5$ ms^{-1} and U_A and U_C. The wind effect is also taken into account in the ~~parameterization~~ parametrization of forced latent and sensitive heat fluxes. Instead of calibrating parameters like the "Stanton" and "Dalton" numbers, respectively for the sensible and latent heat fluxes (Deltares, 2023), we set the model to use C_d also in calculating these heat fluxes.

To account for the uncertainties related to the input wind speed and to calibrate the wind function for the forced heat fluxes, the drag coefficient was adjusted by introducing the corrective parameter α as follows:

$$C_{di} = \alpha C_{di}^0. \tag{B1}$$

Background values of the vertical eddy viscosity and diffusivity (Vicoww, Dicoww) were set as equal to molecular values (i.e. 10^{-6} and 10^{-9} m^{2}s^{-1}, respectively), while all the other calibration parameters (e.g. bottom roughness) were set as default values.

The Secchi ~~disk~~ depth D_s was provided as time series with approximately monthly frequency, as obtained from in-situ measurements. These values were not modified. As objective function for this base calibration, we used the error along the entire water column ϵ_T ε_T.

Code availability. The current version of the model is available at https://github.com/eawag-surface-waters-research/Delft3D/tree/d3d4/ research/surface_heat_transfer under the GPLv3 licence. The exact version of the model used to produce the results used in this paper is archived on Zenodo (Amadori et al., 2024) as are input data and scripts to run the model and produce the plots for all the simulations presented in this paper.

Author contributions. Conceptualization: M.A., M.T.; Data curation: M.A.; Formal analysis: M.A., M.T.; Funding acquisition: D.B., M.T.; Investigation: M.A., A.I.R., D.B., M.T., Methodology: M.A., M.T., Software: M.A., M.T.; Validation: M.A.; Visualization: M.A., M.T.; Writing – original draft preparation: M.A., M.T.; Writing – review & editing M.A., A.I.R., D.B., M.T..

Competing interests. The authors declare no competing interests

Acknowledgements. This work was supported by ESA through the grant AlpLakes (Contract No.4000136401; AO/1-8216/15/I-SBo; Scientific Exploitation of Operational Missions - SEOM S2_4SCI Land and Water Coastal and Inland Waters). MT acknoweldges the support of the project DICAM-EXC (Departments of Excellence 2023–2027, Grant L232/2016), funded by by the Italian Ministry of Universities and Research (MUR). The authors are grateful to Wei Xia for his prompt assistance on the use of DYNO-PODS and to Claudia Giardino, Mariano Bresciani, Sebastiano Piccolroaz, Daniel Odermatt, Fabian Bärenbold and Jorrit Mesman for insightful discussions and literature suggestions, and to the two reviewers for their valuable feedbacks.

References

Amadori, M., Morini, G., Piccolroaz, S., and Toffolon, M.: Involving citizens in hydrodynamic research: A combined local knowledge - numerical experiment on Lake Garda, Italy, Science of The Total Environment, 722, 137 720, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.137720, 2020.

Amadori, M., Giovannini, L., Toffolon, M., van Haren, H., and Dijkstra, H.: Multi-scale evaluation of a 3D lake model forced by an atmospheric model against standard monitoring data, Environmental Modelling and Software, 139, 105 017, https://doi.org/10.1016/j.envsoft.2021.105017, 2021.

Amadori, M., Rahaghi, A. I., Bouffard, D., Toffolon, M., and Runnals, J.: Repository for: "Using automatic calibration to improve the physics behind complex numerical models: An example from a 3D lake model", https://doi.org/10.5281/zenodo.13712738, zenodo Repository, 2024.

Baar, A. W., Boechat Albernaz, M., van Dijk, W. M., and Kleinhans, M. G.: Critical dependence of morphodynamic models of fluvial and tidal systems on empirical downslope sediment transport, Nature Communications, 10, 4903, https://doi.org/10.1038/s41467-019-12753-x, 2019.

Baracchini, T., Chu, P. Y., Šukys, J., Lieberherr, G., Wunderle, S., Wüest, A., and Bouffard, D.: Data assimilation of in situ and satellite remote sensing data to 3D hydrodynamic lake models: a case study using Delft3D-FLOW v4.03 and OpenDA v2.4, Geoscientific Model Development, 13, 1267–1284, https://doi.org/10.5194/gmd-13-1267-2020, 2020a.

Baracchini, T., Hummel, S., Verlaan, M., Cimatoribus, A., Wüest, A., and Bouffard, D.: An automated calibration framework and open source tools for 3D lake hydrodynamic models, Environmental Modelling & Software, 134, 104 787, https://doi.org/https://doi.org/10.1016/j.envsoft.2020.104787, 2020b.

Biemond, B., Amadori, M., Toffolon, M., Piccolroaz, S., van Haren, H., and Dijkstra, H. A.: Deep-mixing and deep-cooling events in Lake Garda: Simulation and mechanisms, Journal of Limnology, 80, https://doi.org/10.4081/jlimnol.2021.2010, 2021.

Blumberg, A. and Mellor, G.: A description of a three-dimensional coastal ocean circulation model, three-dimensional coastal ocean models, Coastal Estuarine Sci., 4, https://doi.org/10.1029/CO004p0001, 1987.

Castelletti, A., Pianosi, F., Soncini-Sessa, R., and Antenucci, J. P.: A multiobjective response surface approach for improved water quality planning in lakes and reservoirs, Water Resources Research, 46, https://doi.org/https://doi.org/10.1029/2009WR008389, 2010.

Chanudet, V., Fabre, V., and van der Kaaij, T.: Application of a three-dimensional hydrodynamic model to the Nam Theun 2 Reservoir (Lao PDR), Journal of Great Lakes Research, 38, 260–269, https://doi.org/https://doi.org/10.1016/j.jglr.2012.01.008, 2012.

Chen, C., Liu, H., and Beardsley, R.: An Unstructured Grid, Finite-Volume, Three-Dimensional, Primitive Equations Ocean Model: Application to Coastal Ocean and Estuaries, Journal of Atmospheric and Oceanic Technology - J ATMOS OCEAN TECHNOL, 20, 159–186, https://doi.org/10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2, 2003.

Cole, T. and Wells, S.: CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.0. User Manual, Instruction Report EL-2000, US Army Engineering and Research Development Center, Vicksburg, 2015.

Dake, J. M. K. and Harleman, D. R. F.: Thermal stratification in lakes: Analytical and laboratory studies, Water Resources Research, 5, 484–495, https://doi.org/https://doi.org/10.1029/WR005i002p00484, 1969.

Deltares: Delft3D-FLOW, User Manual. Version 4.05, Delft Hydraulics, 2023.

Dissanayake, P., Hofmann, H., and Peeters, F.: Comparison of results from two 3D hydrodynamic models with field data: internal seiches and horizontal currents, Inland Waters, 9, 239–260, https://doi.org/10.1080/20442041.2019.1580079, 2019.

El Serafy, G., Gerritsen, H., Hummel, S., Weerts, A., Mynett, A., and Tanaka, M.: Application of data assimilation in portable operational forecasting systems—the DATools assimilation environment, Ocean Dynamics, 57, 485–499, https://doi.org/10.1007/s10236-007-0124-3, 2007.

Fang, X. and Stefan, H. G.: Long-term lake water temperature and ice cover simulations/measurements, Cold Regions Science and Technology, 24, 289–304, https://doi.org/https://doi.org/10.1016/0165-232X(95)00019-8, 1996.

Garcia, M., Ramirez, I., Verlaan, M., and Castillo, J.: Application of a three-dimensional hydrodynamic model for San Quintin Bay, B.C., Mexico. Validation and calibration using OpenDA, Journal of Computational and Applied Mathematics, 273, 428–437, https://doi.org/https://doi.org/10.1016/j.cam.2014.05.003, 2015.

Goudsmit, G.-H., Burchard, H., Peeters, F., and Wüest, A.: Application of k-ϵ turbulence models to enclosed basins: The role of internal seiches, J. Geophys. Res., 107, https://doi.org/10.1029/2001JC000954, 2002.

Henderson-Sellers, B.: Calculating the surface energy balance for lake and reservoir modeling: A review, Reviews of Geophysics, 24, 625–649, 1986.

Hipsey, M. R., Bruce, L. C., Boon, C., et al.: A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON), Geosci. Model Dev., 12, 473–523, https://doi.org/10.5194/gmd-12-473-2019, 2019.

Joehnk, K. and Umlauf, L.: Modelling the metalimnetic oxygen minimum in a medium sized alpine lake, Ecological Modelling, 136, 67–80, https://doi.org/10.1016/S0304-3800(00)00381-1, 2001.

Lesser, G. R., Roelvink, J. v., Van Kester, J., and Stelling, G.: Development and validation of a three-dimensional morphological model, Coastal engineering, 51, 883–915, 2004.

Ma, J., Li, R., Zheng, H., Li, W., Rao, K., Yang, Y., and Wu, B.: Multivariate adaptive regression splines-assisted approximate Bayesian computation for calibration of complex hydrological models, Journal of Hydroinformatics, 26, 503–518, https://doi.org/10.2166/hydro.2024.232, 2024.

Madec, G., Bell, M., Blaker, A., Bricaud, C., Bruciaferri, D., Castrillo, M., Calvert, D., Chanut, J., Clementi, E., Coward, A., Epicoco, I., Éthé, C., Ganderton, J., Harle, J., Hutchinson, K., Iovino, D., Lea, D., Lovato, T., Martin, M., Martin, N., Mele, F., Martins, D., Masson, S., Mathiot, P., Mele, F., Mocavero, S., Müller, S., Nurser, A. G., Paronuzzi, S., Peltier, M., Person, R., Rousset, C., Rynders, S., Samson, G., Téchené, S., Vancoppenolle, M., and Wilson, C.: NEMO Ocean Engine Reference Manual, https://doi.org/10.5281/zenodo.8167700, 2023.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C.: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers, Journal of Geophysical Research: Oceans, 102, 5753–5766, https://doi.org/https://doi.org/10.1029/96JC02775, 1997.

Morel, A. and Antoine, D.: Heating Rate within the Upper Ocean in Relation to its Bio–optical State, Journal of Physical Oceanography, 24, 1652 – 1665, https://doi.org/https://doi.org/10.1175/1520-0485(1994)024<1652:HRWTUO>2.0.CO;2, 1994.

Piccioni, F., Casenave, C., Lemaire, B. J., Le Moigne, P., Dubois, P., and Vinçon-Leite, B.: The thermal response of small and shallow lakes to climate change: new insights from 3D hindcast modelling, Earth System Dynamics, 12, 439–456, https://doi.org/10.5194/esd-12-439-2021, 2021.

Piccolroaz, S., Amadori, M., Toffolon, M., and Dijkstra, H. A.: Importance of planetary rotation for ventilation processes in deep elongated lakes: Evidence from Lake Garda (Italy), Scientific Reports, 9, 2045–2322, https://doi.org/10.1038/s41598-019-44730-1, 2019.

585 Piccolroaz, S., Zhu, S., Ladwig, R., Carrea, L., Oliver, S., Piotrowski, A. P., Ptak, M., Shinohara, R., Sojka, M., Woolway, R. I., and Zhu, D. Z.: Lake Water Temperature Modeling in an Era of Climate Change: Data Sources, Models, and Future Prospects, Reviews of Geophysics, 62, e2023RG000 816, https://doi.org/https://doi.org/10.1029/2023RG000816, 2024.

Poole, H. H. and Atkins, W. R. G.: Photo-electric Measurements of Submarine Illumination throughout the Year, Journal of the Marine Biological Association of the United Kingdom, 16, 297–324, https://doi.org/10.1017/S0025315400029829, 1929.

590 Pothoven, S. A. and Vanderploeg, H. A.: Seasonal patterns for Secchi depth, chlorophyll a, total phosphorus, and nutrient limitation differ between nearshore and offshore in Lake Michigan, Journal of Great Lakes Research, 46, 519–527, https://doi.org/https://doi.org/10.1016/j.jglr.2020.03.013, 2020.

Rahaghi, A. I., Odermatt, D., Anneville, O., Steiner, O. S., Reiss, R. S., Amadori, M., Toffolon, M., Jacquet, S., Harmel, T., Werther, M., Soulignac, F., Dambrine, E., Jézéquel, D., Hatté, C., Tran-Khac, V., Rasconi, S., Rimet, F., and Bouffard, D.: Combined
595 Earth observations reveal the sequence of conditions leading to a large algal bloom in Lake Geneva, Commun Earth Environ, 5, https://doi.org/https://doi.org/10.1038/s43247-024-01351-5, 2024.

Ralston, M. L. and Jennrich, R. I.: Dud, A Derivative-Free Algorithm for Nonlinear Least Squares, Technometrics, 20, 7–14, https://doi.org/10.1080/00401706.1978.10489610, 1978.

Schmid, M. and Köster, O.: Excess warming of a Central European lake driven by solar brightening, Water Resources Research, 52, 8103–
600 8116, https://doi.org/https://doi.org/10.1002/2016WR018651, 2016.

Schmid, M., Hunziker, S., and Wüest, A.: Lake surface temperatures in a changing climate: a global sensitivity analysis, Climatic Change, 124, 301–315, https://doi.org/https://doi.org/10.1007/s10584-014-1087-2, 2014.

Schwindt, S., Callau Medrano, S., Mouris, K., Beckers, F., Haun, S., Nowak, W., Wieprecht, S., and Oladyshkin, S.: Bayesian calibration points to misconceptions in three-dimensional hydrodynamic reservoir modeling, Water Resources Research, 59, e2022WR033 660,
605 https://doi.org/10.1029/2022WR033660, 2022.

Secchi, A.: Relazione delle esperienze fatte a bordo della pontificia pirocorvetta l'Immacolata concezione per determinare la trasparenza del mare; Memoria del P. A. Secchi, Il Nuovo Cimento (1855-1868), 20, 205–238, https://doi.org/10.1007/BF02726911, 1864.

Shatwell, T., Thiery, W., and Kirillin, G.: Future projections of temperature and mixing regime of European temperate lakes, Hydrology and Earth System Sciences, 23, 1533–1551, https://doi.org/10.5194/hess-23-1533-2019, 2019.

610 Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Modelling, 9, 347–404, https://doi.org/https://doi.org/10.1016/j.ocemod.2004.08.002, 2005.

Smagorinsky, J.: "General circulation experiments with the primitive equations", Monthly Weather Review, 91, 99 – 164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2, 1963.

Soulignac, F., Vinçon-Leite, B., Lemaire, B., Martins, J. R., Bonhomme, C., Dubois, P., Mezemate, Y., Tchiguirinskaia, I., Schertzer, D., and
615 Tassin, B.: Performance Assessment of a 3D Hydrodynamic Model Using High Temporal Resolution Measurements in a Shallow Urban Lake, Environmental Modeling & Assessment, 22, 1–14, https://doi.org/10.1007/s10666-017-9548-4, 2017.

Soulignac, F., Danis, P.-A., Bouffard, D., Chanudet, V., Dambrine, E., Guénand, Y., Harmel, T., Ibelings, B. W., Trevisan, D., Uittenbogaard, R., and Anneville, O.: Using 3D modeling and remote sensing capabilities for a better understanding of spatio-temporal heterogeneities of phytoplankton abundance in large lakes, Journal of Great Lakes Research, 44, 756–764,
620 https://doi.org/https://doi.org/10.1016/j.jglr.2018.05.008, 2018.

Stepanenko, V. and Lykossov, V.: Numerical modeling of the heat and moisture transport in a lake-soil system, Russian Meteorology and Hydrology, pp. 69–75, 2005.

Thiery, W., Stepanenko, V. M., Fang, X., Jöhnk, K. D., Li, Z., Martynov, A., Perroud, M., Subin, Z. M., Darchambea, F., Mironov, D., and Lipzig, N. P. M. V.: LakeMIP Kivu: evaluating the representation of a large, deep tropical lake by a set of one-dimensional lake models, Tellus A: Dynamic Meteorology and Oceanography, 66, 21 390, https://doi.org/10.3402/tellusa.v66.21390, 2014.

Tritthart, M., Vanzo, D., Chavarrías, V., Siviglia, A., Sloff, K., and Mosselman, E.: Why do published models for fluvial and estuarine morphodynamics use unrealistic representations of the effects of transverse bed slopes?, Advances in Water Resources, 193, 104 831, https://doi.org/https://doi.org/10.1016/j.advwatres.2024.104831, 2024.

Voudouri, A., Avgoustoglou, E., and Kaufmann, P.: Impacts of Observational Data Assimilation on Operational Forecasts, in: Perspectives on Atmospheric Sciences, edited by Karacostas, T., Bais, A., and Nastos, P. T., pp. 143–149, Springer International Publishing, Cham, 2017.

Wahl, B. and Peeters, F.: Effect of climatic changes on stratification and deep-water renewal in Lake Constance assessed by sensitivity studies with a 3D hydrodynamic model, Limnology and Oceanography, 59, 1035–1052, https://doi.org/10.4319/lo.2014.59.3.1035, 2014.

Wüest, A. and Lorke, A.: Small-scale hydrodynamics in lakes, Annu. Rev. Fluid Mech., 35, 373–412, https://doi.org/10.1146/annurev.fluid.35.101101.161220, 2003.

Xia, W., Shoemaker, C., Akhtar, T., and Nguyen, M.-T.: Efficient Parallel Surrogate Optimization Algorithm and Framework with Application to Parameter Calibration of Computationally Expensive Three-dimensional Hydrodynamic Lake PDE Models, Environ. Model. Softw., 135, 104 910, 2021.

Xia, W., Akhtar, T., and Shoemaker, C. A.: A novel objective function DYNO for automatic multivariable calibration of 3D lake models, Hydrology and Earth System Sciences, 26, 3651–3671, 2022.

Zaneveld, J. and Spinrad, R.: An arc tangent model of irradiance in the sea, Journal of Geophysical Research, 85, https://doi.org/10.1029/JC085iC09p04919, 1980.

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.