Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

A General Lake Model (GLM 2.4) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON)

Matthew R Hipsey[1,*], Louise C Bruce[1], Casper Boon[1], Brendan Busch[1], Cayelan C. Carey[2], David P Hamilton[3], Paul C. Hanson[4], Jordan S. Read[5], Eduardo de Sousa[1], Michael Weber[6], Luke A. Winslow[7]

[1] UWA School of Agriculture & Environment, The University of Western Australia, Crawley WA, 6009, Australia
[2] Department of Biological Sciences, Virginia Tech, Blacksburg VA, USA
[3] Australian Rivers Institute, Griffith University, Nathan QLD, 4111, Australia
[4] Center for Limnology, University of Wisconsin - Madison, Madison WI, USA
[5] Center for Integrated Data Analytics, U.S. Geological Survey, Middleton WI, USA
[6] Department of Lake Research, Helmholtz Centre for Environmental Research - UFZ, Magdeburg, Germany
[7] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy NY, USA

[*] *Correspondence to*: Matthew R. Hipsey (matt.hipsey@uwa.edu.au)

Keywords: lake, stratification, mixing, water balance climate change, water quality, observatory network,
Word count: 12,000 words (text); 18,000 words (text+references+figures&tables+appendix)

Abstract. The General Lake Model (GLM) is a one-dimensional open-source model code designed to simulate the hydrodynamics of lakes, reservoirs and wetlands. GLM was developed to support the science needs of the Global Lake Ecological Observatory Network (GLEON), a network of lake sensors and researchers attempting to understand lake functioning and address questions about how lakes around the world vary in response to climate and land-use change. The scale and diversity of lake types, locations and sizes, as well as the observational data within GLEON, created the need for a robust community model of lake dynamics with sufficient flexibility to accommodate a range of scientific and management needs of the GLEON community. This paper summarises the scientific basis and numerical implementation of the model algorithms, including details of sub-models that simulate surface heat exchange and ice-cover dynamics, vertical mixing and inflow/outflow dynamics. A summary of typical parameter values for lakes and reservoirs collated from a range of sources is included. GLM supports a dynamic coupling with biogeochemical and ecological modelling libraries for integrated simulations of water quality and ecosystem health. An overview of approaches for integration with other models, and utilities for the analysis of model outputs and for undertaking sensitivity and uncertainty assessments is also provided. Finally, we discuss application of the model within a distributed cloud-computing environment, and as a tool to support learning of network participants.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

1 Introduction

Lakes and other lentic (standing) waters support extensive ecosystem services such as water supply, flood mitigation, hydropower, aesthetic and cultural benefits, as well as fisheries and biodiversity (Mueller et al., 2016). Lakes are often considered to be "*sentinels of change*", providing a window into the sustainability of activities in their associated river basins (Williamson et al., 2009). They are also particularly susceptible to impacts from invasive species and land use development, which often lead to water quality deterioration and loss of ecosystem integrity. Recent estimates have demonstrated their significance in the earth system, contributing to heterogeneity in land surface properties and feedbacks to regional and global climate through energy, water and biogeochemical transfers (Martynov et al., 2012, Cole et al., 2007). For example, Tranvik et al. (2009) suggested carbon burial in lakes and reservoirs is substantial on the global scale, on the order of 0.6 Pg yr^{-1}, or four times the oceanic burial rate.

Given the diversity of lakes among continents, region-specific pressures and local management approaches, the Global Lake Ecological Observatory Network (GLEON: gleon.org) was initiated in 2004 as a grass-roots science community with a vision to observe, understand and predict freshwater systems at a global scale (Hanson et al., 2016). In doing so, GLEON has been a leading example of collaborative research within the hydrological and ecological science disciplines. GLEON aims to bring together environmental sensor networks, numerical models, and information technology to explore ecosystem dynamics across a vast range of scales - from an individual lake or reservoir (Hamilton et al., 2015) to regional (Read et al., 2014; Klug et al., 2012), and even global trends (Rigosi et al., 2015; O´Reilly et al., 2015). Ultimately, it is the aim of the network to facilitate primary discovery and synthesis to provide an improved scientific basis for sustainable freshwater resource management.

Environmental modelling forms a critical component of observing systems, as a way to make sense of the "data deluge" (Porter et al., 2012), allowing users to build virtual domains to support knowledge discovery at the system-scale (Ticehurst et al., 2007; Hipsey et al., 2015). In lake ecosystems the tight coupling between physical processes and water quality and ecological dynamics has long been recognised, and models have capitalized on comprehensive understanding of physical processes (e.g., Imberger and Patterson, 1990; Imboden and Wüest, 1995) to use hydrodynamic models as an underpinning basis for coupling to ecological models that have contributed to our understanding of lake dynamics, including aspects such as mixing regimes, eutrophication dynamics (Matzinger et al., 2007), harmful algal bloom dynamics (Chung et al., 2014), and fisheries (Makler-Pick et al., 2009).

In recent decades a range of 1, 2, and 3-dimensional hydrodynamic models has emerged for lake simulation across a diverse range of time scales. Depending on the dimensionality, the horizontal resolution of these models may vary from metres to tens of kilometres, and the spatial resolution from sub-metre to several metres. As in all modelling disciplines, identifying

the most parsimonious model structure and degree of complexity and resolution is challenging, and users in the lake modelling community often tend to rely on heuristic rules or practical reasons for model choice (Mooij et al., 2010). High-resolution models are suited to studying events that occur at the time-scale of flow dynamics, but are not always desirable for ecological studies over longer time scales due to their computational demands and level of over-parameterisation. On the other hand, simple models may be more agile for a particular application, and more suited to parameter identification and scenario testing workflows, but are often less applicable across a wide variety of domains, making them less generalizable.

The lake modelling community has often relied on 1-dimensional (1D) models, which originated to capture lake water balance and thermal stratification dynamics (e.g., Imberger and Patterson, 1981; Peeters et al., 2007; Saloranta and Andersen, 2007; Perroud et al., 2009; Kirillin et al., 2011; Stepanko et al., 2013). Their use is justified given the dominant role of seasonal changes in vertical stratification on lake dynamics, including oxygen dynamics, nutrient and metal cycling and plankton dynamics (Hamilton and Schladow, 1997; Gal et al., 2009). Despite advances in computing power and more readily available 3D hydrodynamic drivers, they continue to remain attractive as they are easily linked with biogeochemical and ecological modelling libraries for complex ecosystem simulations, allowing them to be used to capture the long-term trajectory and resilience of lakes and reservoirs in response to climate change, hydrologic change and land use change; for example, changes to oxygen and nutrient cycles and the increasing risk of algal blooms (e.g., Hu et al., 2016; Snortheim et al., 2017). Furthermore, their low computational requirements relative to 3D models allow for their use in parameter identification routines, making them an attractive balance between process complexity and computational intensity. Nonetheless, there has been a continuing proliferation in the diversity of lake models (Mooij et al., 2010; Janssen et al., 2015), with no clear packages that are suited to the broad range of geographic contexts, time-scales, and science questions and management issues being addressed by the network participants. In acknowledging that there is no single model suitable for all lake applications, a range of open-source community models and tools can enhance scientific capabilities and foster scientific collaboration and combined efforts (Read et al., 2016). To improve scientific collaboration within the limnological modelling community, however, there is an increasing need for a flexible, open-source community model that limnologists can apply to their own lakes (Trolle et al., 2012), as has been common in oceanography, hydrology and climate modelling communities.

In response to this need, the General Lake Model (GLM), a one-dimensional hydrodynamic model for enclosed aquatic ecosystems was developed. The model emerged as a new code from GLEON activities in 2012, and computes the lake water and energy balance by adopting a variable layer structure, allowing for simulation of vertical profiles of temperature, salinity and density, and considering the potential effects of inflows and outflows, surface heating and cooling, mixing and the effect of ice cover on heating and mixing of the lake. GLM is itself a hydrodynamic model, but has dynamic links to biogeochemical models, allowing for exploration of stratification and vertical mixing on the dynamics of biogeochemical cycles, water quality attributes, and lake ecology. The scope and capability of the model has since developed rapidly with

application to numerous lakes within the GLEON network and beyond (e.g., Read et al., 2014; Bueche et al., 2017, Snortheim et al., 2017; Weber et al., 2017; Menció et al., 2017; Bruce et al., 2017). GLM has been designed to be an open-source community model suited to modelling studies across a broad spectrum of lakes, reservoirs and wetlands. It balances complexity of dimensional representation, applicability to a wide range of standing waters, and availability to a broad community (e.g., GLEON). Given that individual applications of the model are not able to describe the full array of features and details of the model structure, the aim of this paper is to present a complete description of GLM, including the scientific background (Section 2), model code organization (Section 3), approach to coupling with biogeochemical models (Section 4), and to overview use of the model within the context of GLEON specific requirements for model analysis, integration and education (Section 5-6).

2 Model Overview

2.1 Background and layer structure

GLM adopts a 1D approach for simulating lake mixing processes by resolving a vertical series of layers that describe the variation in water column properties. Users may configure any number of inflows and outflows, and more advanced options exist for simulating aspects of the water and heat balance. Depending on the context of the simulation, either daily or hourly meteorological time-series data for surface forcing is required, and daily time-series of volumetric inflow and outflow rates can also be supplied. The model is suitable for operation in a wide range of climate conditions and is able to simulate ice formation, as well as accommodating a range of atmospheric forcing conditions.

Although GLM is a new model code written in the C programming language, the core layer structure and mixing algorithms is founded on principles and experience from model platforms including the Dynamic Reservoir Simulation Model (DYRESM; Imberger and Patterson, 1981; Hamilton and Schladow, 1997) and the Dynamic Lake Model (DLM; Chung et al., 2008). Other variations have been introduced to extend this underlying approach through applications to a variety of lake and reservoir environments, to which the reader is also referred (e.g., Hocking & Patterson, 1991; McCord & Schladow, 1998; Gal et al., 2003; Yeates and Imberger, 2003). The layer structure is numbered from the lake bottom to the surface, and adopts the flexible Lagrangian layer scheme first introduced by Imberger et al. (1978) and Imberger & Patterson (1981). The approach defines each layer as a 'control volume' that can change thickness by contracting and expanding in response to inflows, outflows, mixing with adjacent layers, and surface mass fluxes. Layer thickness limits are enforced to adequately resolve the vertical density gradient with fine resolution occurring in the metalimnion and thicker cells where mixing is occurring, as depicted schematically in Figure 1. Unlike fixed-grid (Eulerian) design of most 1D lake and ocean models, where mixing algorithms are typically based on resolving vertical velocities, it has been reported that numerical diffusion at the thermocline can be restricted by this approach (depending on the user-defined minimum (h_{min}) and maximum (h_{max}) layer thickness limits set by the user), making it particularly suited to long-term investigations, and requiring limited site-

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

specific calibration (Patterson et al., 1984; Hamilton & Schladow, 1997; Bruce et al., 2017). Layers each have a unique density computed based on the local salinity and temperature, and when sufficient energy becomes available to overcome density instabilities between adjacent layers, they will merge, thereby accounting for the process of mixing. For deeper systems, a stable vertical density gradient will form in response to periods of high solar radiation creating warm, less-dense conditions near the surface with cooler conditions deeper in the water, separated by a metalimnion region which includes the thermocline. The number of layers, $N_{LEV}(t)$, is adjusted throughout the simulation to maintain homogenous properties within a layer. Initially, the layers are assumed to be of equal thickness, and the initial number of layers, $N_{LEV}(t = 0)$. As the model simulation progresses, density changes due to surface heating, vertical mixing, and inflows and outflows lead to dynamic changes in the layer structure, associated with layers amalgamating, expanding, contracting or splitting.

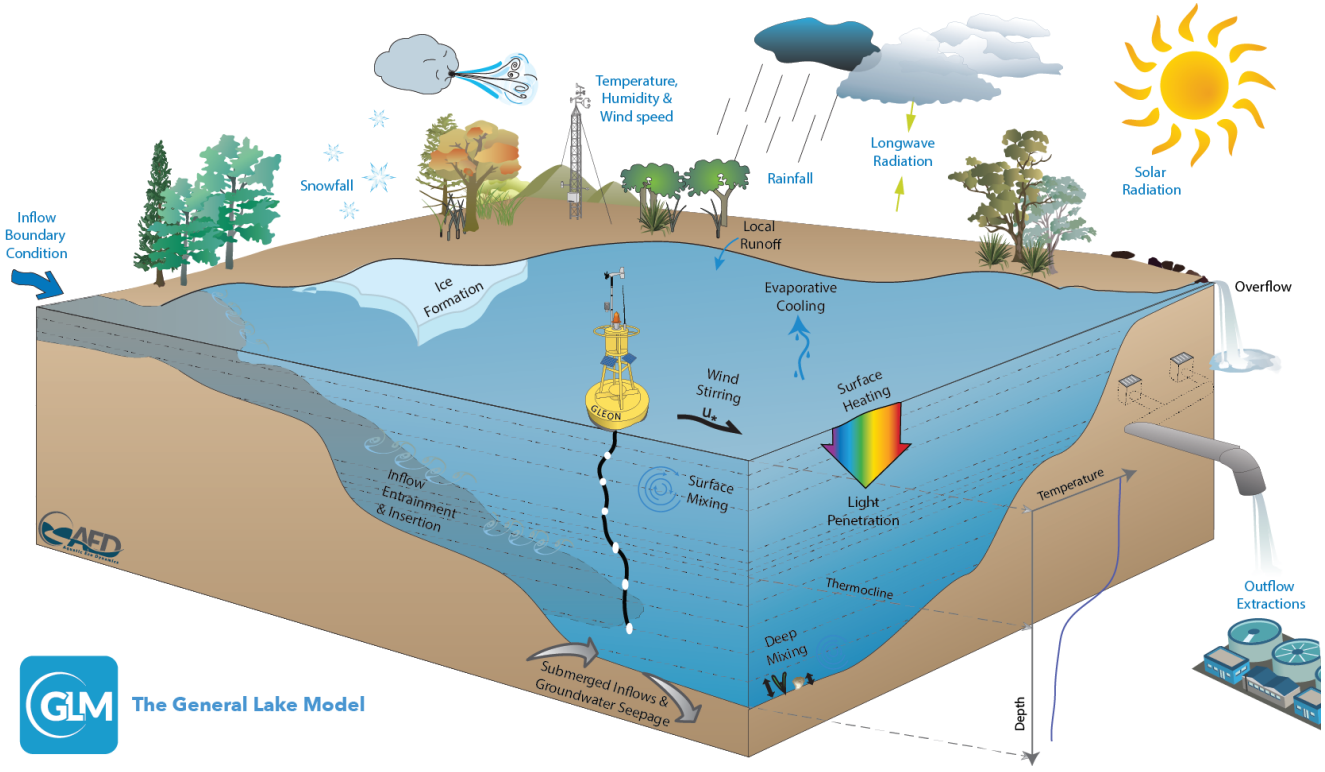

Figure 1: Schematic of a GLM simulation domain, input information (blue text) and key simulated processes (black text).

As layers change, their volumes change based on the site-specific hypsographic curve, whereby the overall lake volume is defined as $\int A(h)dh$, and the elevation (h), and area (A) relationship must be provided as a series of points based on bathymetric data. Layer volumes are determined by interpolating layer area at the appropriate height in the lake basin,

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

whereby $A_i = f(h_i)$, and i is the layer number. This computation requires the user provides a number N_{BSN} of depths with corresponding areas, and the volumes are estimated as:

$$V_b = V_{b-1} + [A_{b-1} + 0.5(A_b - A_{b-1})](h_b - h_{b-1}) \quad (1)$$

where $1 < b \leq N_{BSN}$. Using the raw hyposgraphic data, a refined depth-area-volume relationship is internally computed using finer depth increments (e.g., ~ 0.1 m), giving N_{MORPH} levels that are used for subsequent calculations. The area and volume at the depth of each increment, h_z is interpolated from the supplied information as:

$$V_z = V_{b-1}\left(\frac{h_z}{h_{b-1}}\right)^{\alpha_b} \quad \text{and} \quad A_z = A_{b-1}\left(\frac{h_z}{h_{b-1}}\right)^{\beta_b} \quad (2)$$

where V_z and A_z are the volume and area at each of the elevations of the refined depth vector, and V_z refers to the nearest b level below h_z such that $h_{b-1} < h_z$. The interpolation coefficients are computed as:

$$\alpha_b = \left[\frac{log_{10}\left(\frac{V_{b+1}}{V_b}\right)}{log_{10}\left(\frac{h_{b+1}}{h_b}\right)}\right] \quad \text{and} \quad \beta_b = \left[\frac{log_{10}\left(\frac{A_{b+1}}{A_b}\right)}{log_{10}\left(\frac{h_{b+1}}{h_b}\right)}\right]. \quad (3)$$

The density in each layer i is computed based on the temperature, T, and salinity, S, at any given time according to the UNESCO (1981) equation of state whereby $\rho_i = \rho(T_i, S_i)$. Density calculations can also be customised as required.

Because this approach assumes layer properties are laterally averaged, the model is suitable for investigations where resolving the horizontal variability is not a requirement of the study. This is often the case for ecologists and biogeochemists studying natural lakes (e.g., Gal et al., 2009), managers simulating drinking water reservoirs (e.g., Weber et al., 2017), or mining pit lakes (e.g., Salmon et al., 2017), or for analyses exploring the coupling between lakes and regional climate (e.g., Stepanenko et al., 2013). Further, whilst the model is able to resolve vertical stratification, it may also be used to simulate shallow lakes, wetlands, wastewater ponds and other small waterbodies that experience well-mixed conditions. In this case, the layer resolution, with upper and lower layer bounds specified by the user, will automatically simplify, and mass and energy will continue to be conserved.

2.2 Water balance

The model solves the water balance of the lake domain by including several user-configurable water fluxes. The components include surface mass fluxes (evaporation, rainfall and snowfall), inflows (surface inflows, submerged inflows and local runoff from the surrounding exposed lake bed area) and outflows (withdrawals, overflow and seepage). The dynamics of inflows and outflows modify the overall lake water balance on a daily time-step, and may impact upon the layer structure by adding, merging or removing layers (described in Sect. 2.6). In addition, the mass balance the surface layer is computed at each model time step (usually hourly), by modifying the surface layer height according to:

$$\frac{dh_S}{dt} = E + S + f_R R + Q_R/A_s \quad (4)$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

where h_S is the top height of the surface layer (m), t is the time (days), E is the evaporation mass flux computed from the heat flux ϕ_E (W m^{-2}) described below, R is rainfall and S is snowfall (m day^{-1}), and f_R is a user-definable scaling factor that may be applied to increase or reduce the rainfall data (default = 1). Q_R is an optional term to account for runoff to the lake from the exposed banks, which may be important in reservoirs with a large drawdown range, or wetlands where periodic drying of the lake may occur. The runoff volume generated is averaged across the current lake area (A_s), and the amount is calculated using a simple model based on exceedance of a threshold rainfall intensity, R_L (m day^{-1}), and runoff coefficient:

$$Q_R = f_{ro}(f_R R - R_L)(A_{max} - A_s) \quad (5)$$

where f_{ro} is the runoff coefficient, defined as the fraction of rainfall that is converted to runoff at the lake's edge, and A_{max} is the maximum possible area of inundation of the lake (as defined by the area provided by the user at N_{BSN} area value).

Note mixing dynamics (i.e. the merging or splitting of layers to enforce the layer thickness limits), will impact the thickness of the surface mixed layer, z_{SML}, but not change the overall lake height. However, in addition to the terms in Eq. 4, h_S will also be modified as a result of ice formation/melt, and river inflows, withdrawals, seepage or overflows impacting upon the surface layer, which are described in subsequent sections; these are in addition to the above described terms.

2.3 Surface energy balance

A balance of shortwave and longwave radiation fluxes, and sensible and evaporative heat fluxes determine the net cooling and heating for GLM. The general heat budget equation is described as:

$$\left[\frac{c_p}{A_s z_{sml}}\right]\frac{dT_S}{dt} = \phi_{SW_S} - \phi_E + \phi_H + \phi_{LWin} - \phi_{LWout} \quad (6)$$

where c_p is the specific heat capacity of water (4186 J kg^{-1} °C^{-1}), T_S is the surface temperature of the surface mixed layer and z_{sml} is the depth of the surface mixed layer. The RHS heat flux terms, including several options for customizing the individual surface heat flux components, are expanded upon individually below.

2.3.1 Solar heating and light penetration

Solar radiation is the key driver of the lake thermodynamics, however, data may not always be available from a nearby pyranometer. Users may choose to either have GLM compute surface irradiance from a theoretical approximation based on the Bird Clear Sky insolation model (BCSM) (Bird, 1984), modified for cloud cover and latitude, or alternatively, hourly or daily solar radiation intensity data may be specified directly. If the BCSM is used, then $\hat{\phi}_{SW}$ is calculated from (Bird, 1984; Luo et al., 2010):

$$\hat{\phi}_{SW} = \frac{\hat{\phi}_{DB} + \hat{\phi}_{AS}}{1 - (\alpha_{SW}\,\alpha_{SKY})} f(C) \quad (7)$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

where the model computes total irradiance, $\hat{\phi}_{SW}$ (W m^{-2}) from direct beam $\hat{\phi}_{DB}$, and atmospheric scattering $\hat{\phi}_{AS}$ components (refer to Appendix A for a detailed outline of the BCSM equations and parameters). In GLM, the clear sky value is reduced according to the cloud cover, C, according to:

$$f(C) = 0.66182\,C^2 - 1.5236\,C + 0.98475 \quad (8)$$

which is based on a polynomial regression of cloud data from Perth Airport, Australia, compared against nearby sensor data ($R^2 = 0.952$; see also Luo et al., 2010).

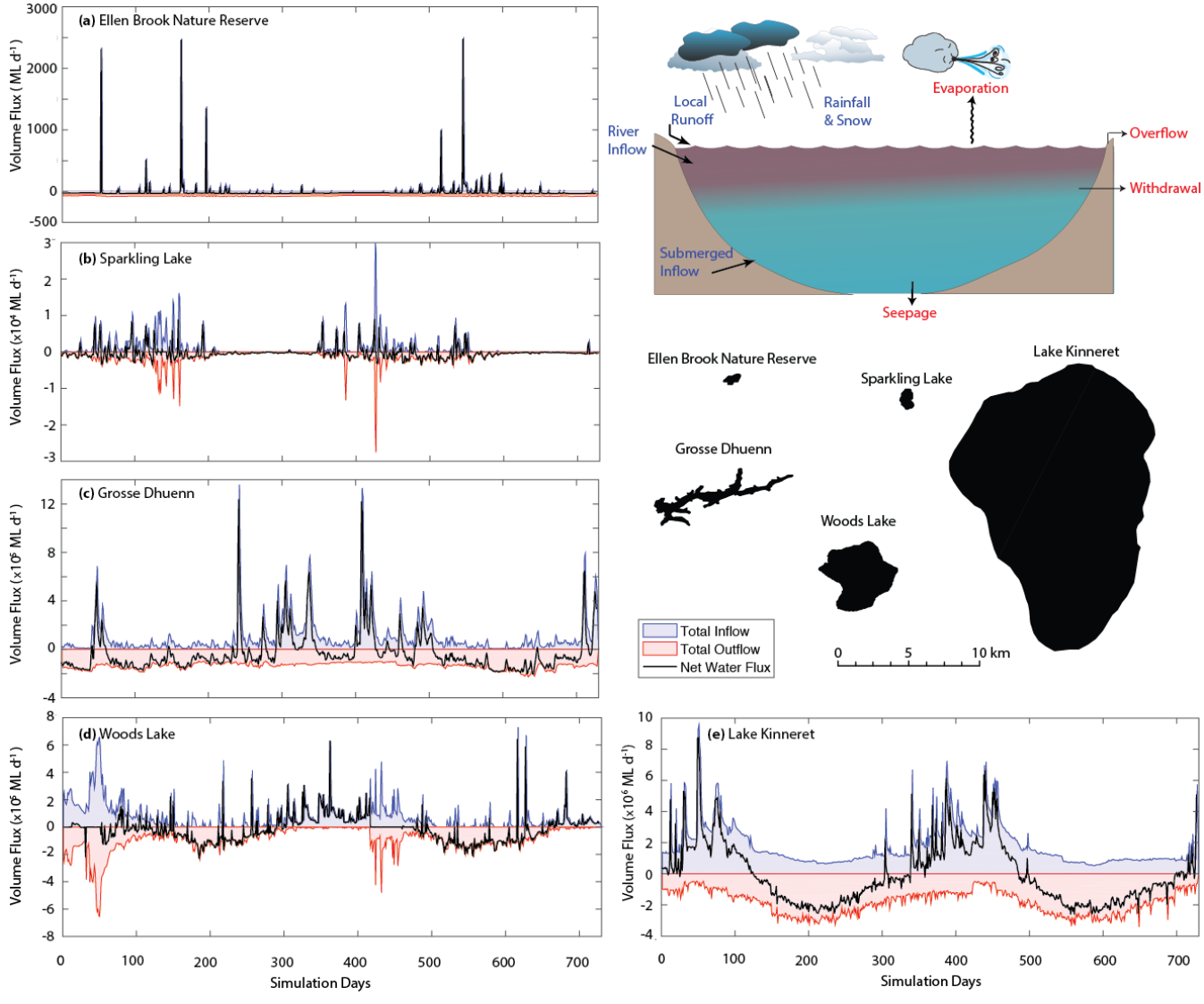

Figure 2: A two-year times-series of the simulated daily water balance for five example lakes, a-e, that range in size and hydrology. The water balance components summarised are depicted schematically in the inset. For more information about each lake and the simulation configuration refer to the Data availability section.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

The albedo, α_{SW}, is the reflected fraction of $\hat{\phi}_{SW}$, with several computation options based on the angle of incident radiation, selected via the `radmode` option in the model configuration file:

Option 1 : Daily approximation, Hamilton & Schladow (1997)

$$\alpha_{SW} = \begin{cases} 0.08 + 0.02\ sin\left[\frac{2\pi}{365}d - \frac{\pi}{2}\right] & \text{:northern hemisphere} \\ 0.08 & \text{:equator} \\ 0.08 - 0.02\ sin\left[\frac{2\pi}{365}d - \frac{\pi}{2}\right] & \text{:southern hemisphere} \end{cases} \quad (9a)$$

Option 2 : Briegleb et al. (1986)

$$\alpha_{SW} = \frac{1}{100}\left(\frac{2.6}{1.1cos(\Phi_{zen})^{1.7}+0.065} + 15[cos(\Phi_{zen}) - 0.1][cos(\Phi_{zen}) - 0.5][cos(\Phi_{zen}) - 1]\right) \quad (9b)$$

Option 3 : Yajima and Yamamoto (2014)

$$\alpha_{SW} = 0.001\ \mathrm{RH}\ [\cos(\Phi_{zen})]^{0.33} - 0.001\ U_x\ [\cos(\Phi_{zen})]^{-0.57} - 0.001\ \varsigma\ [\cos(\Phi_{zen})]^{0.829} \quad (9c)$$

where d is the day of the year, and Φ_{zen} is the solar zenith angle (radians) as outlined in Appendix A, RH is the relative humidity, ς is the atmospheric diffuse radiation, d is the day of year, and U_x is wind speed. The second (oceanic) and third (lacustrine) options allow for diel and seasonal variation of albedo from approximately 0.01 to 0.4 depending on the sun-angle (Figure 3).

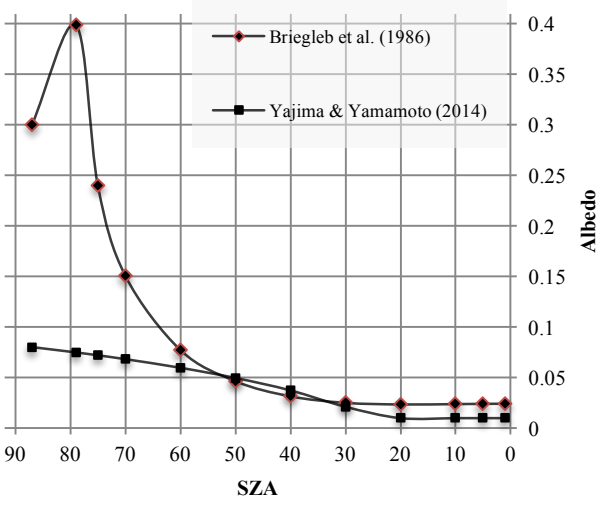

Figure 3: Variation of albedo (α_{SW}) with solar zenith angle (SZA = $2\pi\ \Phi_{zen}/180$, degrees) for `radmode` 2 and 3.

Shortwave radiation penetration into the lake and through the layers is modelled according to the Beer-Lambert Law:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

$$\phi_{SW}(z) = (1 - \alpha_{SW})\, f_{SW} f_{PAR}\, \hat{\phi}_{SW}\, exp[-K_w z] \tag{10}$$

where z is the depth of the layer from the surface, f_{SW} is a scaling factor that may be applied and adjusted as part of the calibration process, and K_w is the light extinction coefficient (m^{-1}). K_w may be set by the user as constant or linked to the water quality model (e.g. FABM or AED2, see Sect 4) in which case the extinction coefficient will change as a function of depth and time according to the concentration of dissolved and particulate constituents. Beer's Law is only applied for the photosynthetically active fraction (PAR) component, f_{PAR}, which is set as 45% of the incident light. The amount of light heating the surface layer, ϕ_{SW_S}, is therefore the above photosynthetically average fraction that is attenuated across z_{SML}, plus the remaining $(1 - f_{PAR})$ fraction which accounts for near infra-red and ultraviolet bandwidths of the incident shortwave radiation with significantly higher attenuation coefficients (Kirk, 1994).

In some applications, the extent to which the benthos has a suitable light climate is a good indicator of benthic productivity, and a proxy for the type of benthic habitat that might emerge. In addition to the light profiles, GLM predicts the benthic area of the lake where light intensity exceeds a user defined value (Figure 4), $\phi_{BEN_{crit}}$.

$$A_{BEN} = A_s - A(h_{BEN}) \tag{11}$$

where $h_{BEN} = h_{SURF} - z_{BEN}$, and z_{BEN} is calculated from Beer's law:

$$z_{BEN} = ln\left[\frac{\phi_{BEN_{crit}}}{\phi_{SW_S}}\right]\frac{-1}{K_w} \tag{12}$$

The daily average benthic area above the threshold is reported in the `lake.csv` summary file as a percentage (A_{BEN}/A_S).

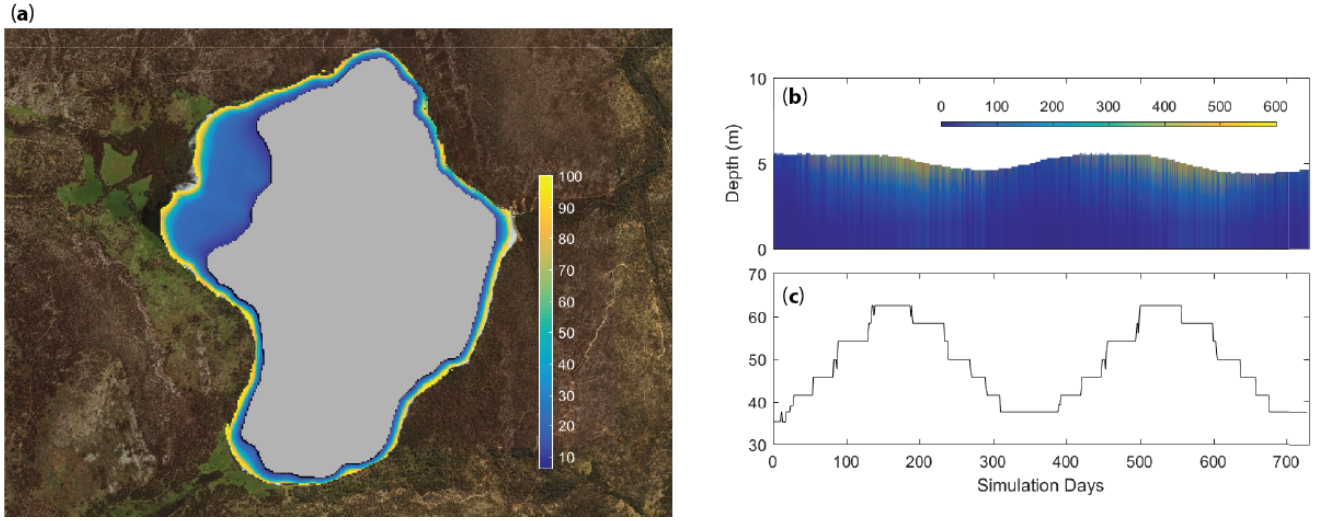

Figure 4: Example light data outputs from a GLM application to Woods Lake, Australia, showing a) the ratio of benthic to surface light, $\phi_{SW_{BEN}}/\phi_{SW_S}$ (%), overlain on the lake map based on the bathymetry, b) a time-series of the depth variation in light (W m^{-2}), and c) a time-series of A_{BEN}/A_S (%).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

2.3.2 Longwave radiation

Longwave radiation can either be specified as net flux, incoming flux or, if there is no radiation data from which longwave radiation can be computed, then it may be calculated by the model internally based on the cloud cover fraction and air temperature. Net longwave radiation is described as:

$$\phi_{LW_{net}} = \phi_{LW_{in}} - \phi_{LW_{out}} \tag{13}$$

where

$$\phi_{LW_{out}} = \varepsilon_w \sigma [T_s + 273.15]^4 \tag{14}$$

and σ is the Stefan-Boltzman constant and ε_w the emissivity of the water surface, assumed to be 0.985. If the net or incoming longwave flux is not provided, the model will compute the incoming flux from:

$$\phi_{LW_{in}} = (1 - \alpha_{LW})\, \varepsilon_a^*\, \sigma\, [T_a + 273.15]^4 \tag{15}$$

where α_{LW} is the longwave albedo (0.03), and the emissivity of the atmosphere is computed considering emissivity of cloud-free conditions (ε_a), based on air temperature (T_a) and vapour pressure, extended to account for reflection from clouds, such that $\varepsilon_a^* = f(T_a, C)$ calculated from (Henderson-Sellers, 1986):

$$\varepsilon_a^* = \begin{cases} (1 + 0.275C)\left(1 - 0.261 \exp\left[-0.000777\, {T_a}^2\right]\right) & \text{Option 1: Idso and Jackson (1969)} \\ (1 + 0.17\, C^2)\, (9.365 \times 10^{-6} [T_a + 273.15]^2) & \text{Option 2: Swinbank (1963)} \\ (1 + 0.275\, C)\, 0.642 \left(\frac{e_a}{T_a}\right)^{\frac{1}{7}} & \text{Option 3: Brutseart (1975)} \\ \left[(1 - C^{2.796})\, 1.24 \left(\frac{e_a}{T_a}\right)^{\frac{1}{7}} + 0.955\, C^{2.796}\right] & \text{Option 4: Yajima and Yamamoto (2014)} \end{cases} \tag{16a-d}$$

where, C is the cloud cover fraction (0-1), e_a the air vapour pressure calculated from relative humidity, and options 1-4 are chosen via the `cloudmode` variable. Note that cloud cover is typically reported in octals (1-8) with each value depicting a fraction of 8. So a value of 1 would correspond to a fraction of 0.125. Some data may also include cloud type and their respective heights. If this is the case, good results have been reported by averaging the octal values for all cloud types to get an average value of cloud cover

If longwave radiation data does not exist and cloud data is also not available, but solar irradiance is measured, then it is possible to get GLM to compare the measured and theoretical (BCSM) solar irradiance to approximate the cloud cover fraction. This option utilises the above relation in Eq. 7 to compute $\hat{\phi}_{SW}$, and clouds are approximated by assuming that $\hat{\phi}_{SW_{OBS}} / \hat{\phi}_{SW_{BCSM}} = f(C)$. Note that if neither shortwave or longwave radiation are provided, then the model will use the BCSM to compute incoming solar irradiance and cloud cover will be assumed to be 0.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

2.3.3 Sensible and latent heat transfer

The model accounts for the surface fluxes of sensible heat and latent heat using commonly adopted bulk aerodynamic formulae. For sensible heat:

$$\phi_H = -\rho_a c_p C_H U_x (T_s - T_a) \tag{17}$$

where c_p is the specific heat capacity of air (1005 J kg^{-1} °C^{-1}), C_H is the bulk aerodynamic coefficient for sensible hear transfer (~1.3x10^{-3}), T_a the air temperature (°C) and T_s the temperature of the surface layer (°C). The air density is in kg m^{-3} and computed from $\rho_a = 0.348\,(1 + r)/(1 + 1.61r)\,p/T_a$, where p is air pressure (hPa) and r is the mixing ratio, which is used to compute the gas constant.

For latent heat:

$$\phi_E = -\rho_a C_E\, \lambda\, U_x \frac{\kappa}{p}\,(e_s[T_s] - e_a[T_a]) \tag{18}$$

where C_E is the bulk aerodynamic coefficient for latent heat transfer, e_a the air vapour pressure, e_s the saturation vapour pressure (hPa) at the surface layer temperature (°C), κ is the ratio of molecular weight of water to molecular weight of air (= 0.622) and λ is the latent heat of vaporisation. The vapour pressure can be calculated by the following formulae:

$$e_s[T_s] = exp\left[2.3026\left(7.5\frac{T_s}{T_s+237.3}\right) + 0.7858\right] \quad \text{Option 1 : TVA (1972) - Magnus-Tetens} \tag{19a}$$

$$e_s[T_s] = exp\left[6.1094\left(\frac{17.625\,T_s}{T_s+243.04}\right)\right] \quad \text{Option 2 : August-Roche-Magnus} \tag{19b}$$

$$e_s[T_s] = 10^{\left(9.28603523\frac{2322.37885\,T_s}{T_s+273.15}\right)} \quad \text{Option 3 : Tabata (1973) - Linear} \tag{19c}$$

$$e_a[T_a] = \frac{RH}{100} e_s[T_a] \tag{20}$$

Correction for non-neutral atmospheric stability : For long-time integrations (i.e., seasonal), the bulk-transfer coefficients for momentum, C_D, sensible heat, C_H, and latent heat, C_E, can be assumed approximately constant because of the negative feedback between surface forcing and the temperature response of the water body (e.g. Strub and Powell, 1987). At finer timescales (hours to weeks), the thermal inertia of the water body is too great and so the transfer coefficients must be specified as a function of the degree of atmospheric stratification experienced in the internal boundary layer that develops over the water (Woolway et al. 2017). Monin and Obukhov (1954) parameterised the stratification in the air column using the now well-known stability parameter, z/L, which is used to define corrections to the bulk aerodynamic coefficients C_H and C_E, using the numerical scheme presented in Appendix B. The corrections may be applied as options in the model, and

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

requires that the measurement of wind speed, air temperature and relative humidity within the internal boundary layer over the lake surface are supplied at an hourly resolution.

Still-air limit : The above formulations only apply when sufficient wind exists to create a defined boundary layer over the surface of the water. As the wind tends to zero (the 'still-air limit'), Eqs. 16-17 are no longer appropriate as they do not account for free convection directly from the water surface. This is a relatively important phenomenon for small lakes, cooling ponds and wetlands since they tend have small fetches that limit the build-up of wind speed. These water bodies are often sheltered from the wind and will develop surface temperatures warmer than the atmosphere for considerable periods. Therefore, we optionally augment Eqs. 16-17 with calculations under low wind-speed conditions by calculating the evaporative and sensible heat flux values for both the given U_x and for an assumed $U_x = 0$. The chosen value is found by taking the maximum value of the two calculations:

$$\phi_{H,E}^{*} = \max\left(\phi_{E,H}, \phi_{E,H_0}\right) \tag{21}$$

where ϕ_0 is the zero-wind flux, given below, and applies for both evaporative and sensible heat fluxes, and $\phi_{E,H}$ is calculated from Eqs. 16-17. The two zero-wind speed heat flux equations are taken from TVA (1972), but modified slightly to return energy flux in SI units (W m^{-2}):

$$\phi_{E_0} = \rho_s \lambda \alpha_e (C_0 - C_a) \tag{22}$$

$$\phi_{H_0} = \alpha_h (T_s - T_a)$$

$$\alpha_e = 2.283\times10^{-3}\, \xi \frac{v}{c_p \rho_s}\left[g\frac{|\rho_a - \rho_o|}{\rho_a \nu\, a}\right]^{1/3} \tag{23}$$

$$\alpha_h = 2.283\times10^{-3}\, \xi\, v\left[g\frac{|\rho_a - \rho_o|}{\rho_a \nu\, a}\right]^{1/3}$$

where $C = \kappa\, e/p$, with the appropriate vapour pressure values, e, for both surface and ambient atmospheric values. Here, v is the molecular heat conductivity of air (0.1 kJ m^{-1} h^{-1} K^{-1}), ν is the kinematic viscosity of the air (0.0548 m^2 h^{-1}], ρ_o is the density of the saturated air at the water surface temperature, ρ_s is the density of the surface water, ξ is a roughness correction coefficient for the lake surface (0.5) and a is the molecular heat diffusivity of air (0.077 m^2 h^{-1}). Note that the impact of low wind speeds on the drag coefficient is captured by the modified Charnock relation (Eq. A24), which includes an additional term for the smooth flow transition (see also Figure A1).

Wind sheltering: Wind-sheltering maybe parameterised according to several methods based on the context of the simulation and data available. Hipsey et al. (2003) presented a simple adjustment to the bulk transfer equation to account for the effect of wind-sheltering in small reservoirs. The method employs the use of a shelter index by accounting for the length scale associated with the vertical obstacle relative to the horizontal length scale associated with the water body itself. Within

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

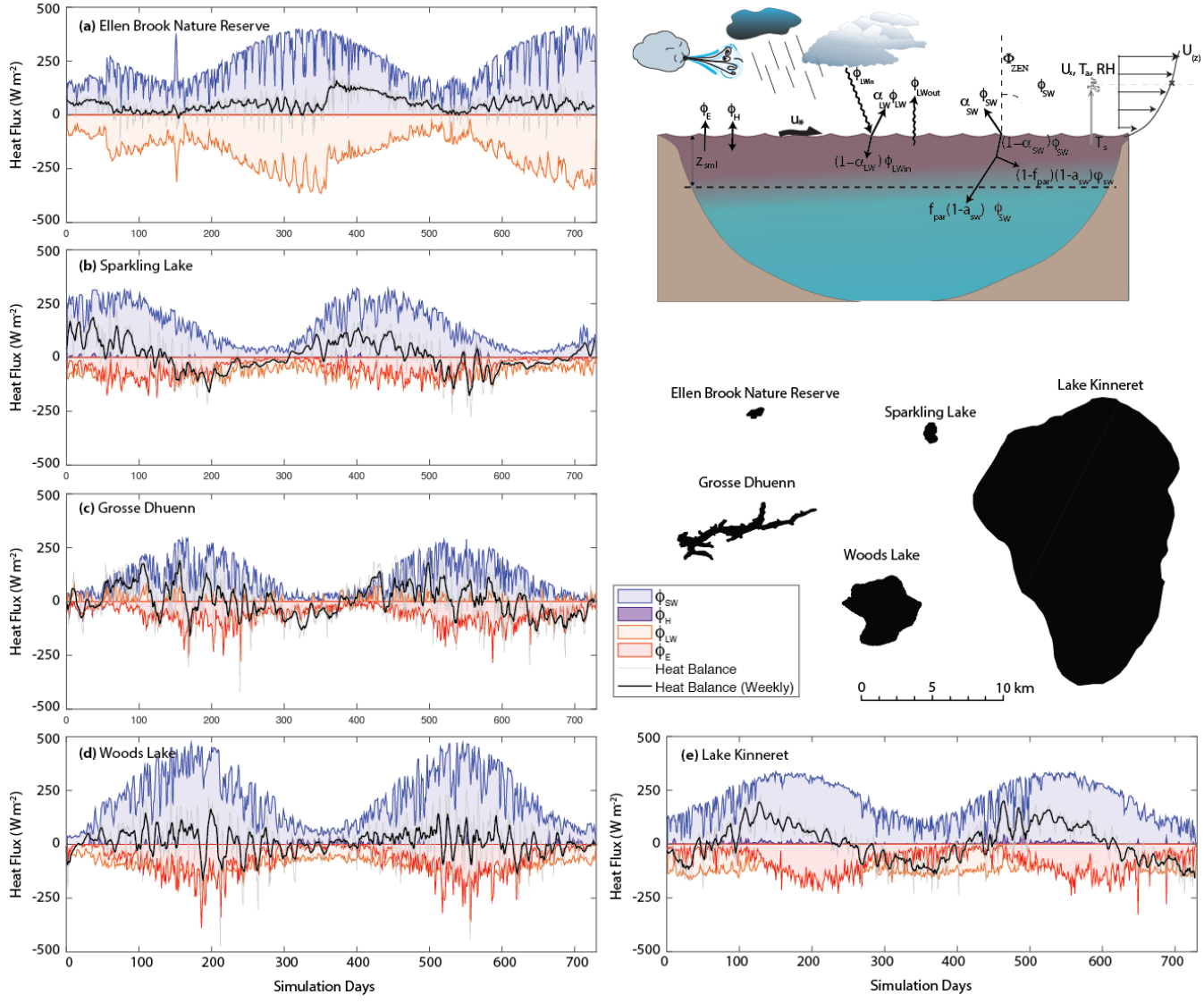

Figure 5: A two-year times-series of the simulated daily heat fluxes for five example lakes, a-e, that were depicted in Figure 2. The heat balance components summarised are depicted schematically in the inset, as described in Sect. 2.3.

GLM, users may specify the degree of sheltering or fetch limitation by optionally supplying the model with the wind direction, and a table linking direction and a wind scaling factor. Alternatively, if the direction-specific data is not available, an effective wind-sheltering coefficient has been implemented that reduces the effective surface area for heat and momentum fluxes:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

$$A_E = \begin{cases} A_S \tanh\left(\frac{A_S}{A_C}\right) & \text{Yeates \& Imberger (2003)} \quad (24a) \\ \frac{D^2}{2} cos^{-1}\left(\frac{x_\tau}{D}\right) - \frac{x_\tau}{D}\sqrt{D^2 - {x_\tau}^2} & \text{Markfort et al. (2009)} \quad (24b) \end{cases}$$

where A_C is the critical area. In GLM, the ratio of the effective area to the total area of the lake A_E/A_S is then used to scale U_x as a means of capturing the average wind speed over the entire lake surface.

2.4 Snow and ice dynamics

The algorithms for GLM ice and snow dynamics are based on previous ice modelling studies (Patterson and Hamblin, 1988; Gu and Stefan, 1993; Rogers et al., 1995; Vavrus et al., 1996; Launiainen and Cheng, 1998; Magee et al., 2016). To solve the heat transfer equation, the ice model uses a quasi-steady assumption that the time scale for heat conduction through the ice is short relative to the time scale of meteorological forcing (Patterson and Hamblin, 1988; Rogers et al., 1995). The steady-state conduction equations, which allocate shortwave radiation into two components, a visible (A1=70%) and an infra-red (A2=30%) spectral band, are used with a three-component ice model that includes blue ice (or black ice), white ice (or snow ice) and snow (see Eq. 1 and Fig. 5 of Rogers et al., 1995). White ice is generated in response to flooding, when the mass of snow that can be supported by the buoyancy of the ice cover is exceeded (see Eq. 13 of Rogers et al., 1995). By assigning appropriate boundary conditions to the interfaces and solving the quasi-steady state of heat transfer numerically, the model computes the upward conductive heat flux from the ice or snow cover to the atmosphere, ϕ_0. The estimation of ϕ_0 involves the application of an empirical equation (Ashton, 1986) to estimate snow conductivity (K_s) from its density (Figure 6).

At the ice (or snow) surface, a heat flux balance is employed to provide the condition for surface melting:

$$\begin{aligned} \phi_0(T_0) + \phi_{net}(T_0) &= 0 \qquad T_0 < T_m \\ \phi_{net}(T_0) &= -\rho L \frac{dh_i}{dt} \qquad T_0 = T_m \end{aligned} \quad (25)$$

where L is the latent heat of fusion (see physical constants, Table 1), h_i is the height of the upper snow or ice layer, t is time, ρ is the density of the snow or ice, determined from the surface medium properties, T_0 is the temperature at the solid surface, T_m is the melt-water temperature (0°C) and $\phi_{net}(T_0)$ is the net incoming heat flux, at the solid surface:

$$\phi_{net}(T_0) = \phi_{LWin} - \phi_{LWout}(T_0) + \phi_H(T_0) + \phi_E(T_0) + \phi_R(T_0) \quad (26)$$

where ϕ_{LWin} and ϕ_{LWout} are incoming and outgoing longwave radiation, ϕ_H and ϕ_E are sensible and evaporative heat fluxes between the solid boundary and the atmosphere, and ϕ_R is the heat flux due to rainfall. These heat fluxes are calculated as above with modification for determination of vapor pressure over ice or snow (Gill, 1982) and the addition of the rainfall heat flux (Rogers et al., 1995). T_0 is determined using a bilinear iteration until surface heat fluxes are balanced (i.e. $\phi_0(T_0) = -\phi_{net}(T_0)$) and T_0 is stable (± 0.001°C). In the presence of ice (or snow) cover, surface temperature $T_0 > T_m$ indicates that

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

energy is available for melting. The amount of energy for melting is calculated by setting $T_0 = T_m$ to determine the reduced thickness of snow or ice (as shown in Eq. 25).

Accretion or ablation of ice is determined through the heat flux at the ice-water interface, q_f. Solving for heat conduction through ice yields:

$$q_f = q_0 - A_1\hat{\phi}_{SW}(1 - \exp[-K_{s1}h_{snow} - K_{w1}h_{white} - K_{b1}h_{blue}]) - A_2\hat{\phi}_{SW}(1 - \exp[-K_{s2}h_{snow} - Kh_{white} - K_{b2}h_{blue}]) - Q_{white}h_{snow} \quad (27)$$

where $\hat{\phi}_{SW}$ is the shortwave radiation penetrating the surface, K refers to the light attenuation coefficient of the ice and snow components designated with subscripts *s, w* and *e* for snow, blue ice and snow ice respectively, and h refers to the thickness of snow, white ice (snow ice) and blue ice. Q_{white} is a volumetric heat flux for formation of snow ice, which is given in Eq. 14 of Rogers et al. (1995). Ice and snow light attenuation coefficients in GLM are fixed to the same values as those given by Rogers et al. (1995). Shortwave albedo for the ice or snow surface is a function of surface medium (snow or ice), surface temperature and ice or snow thickness (see Table 2, Vavrus et al., 1996). Values of albedo derived from these functions vary from 0.08 to 0.6 for ice and from 0.08 to 0.7 for snow.

The imbalance between q_f and the heat flux from the water to the ice, q_w, gives the rate of change of ice thickness at the interface with water:

$$\frac{dh_{blue}}{dt} = \frac{q_f - q_w}{\rho_{blue}L} \quad (28)$$

where ρ_{blue} is the density of blue ice and q_w is given by a finite difference approximation of the conductive heat flux from water to ice:

$$q_w = -K_m\frac{\Delta T}{\Delta z}, \quad (29)$$

where K_m is molecular conductivity and ΔT is the temperature difference between the surface water and the bottom of the ice, which occurs across an assigned depth Δz. A value for Δz of 0.5 m is usual, based on the reasoning given in Rogers *et al.* (1995) and the typical vertical water layer resolution of a model simulation (0.125 – 1.5 m). Note that a wide variation in techniques and values is used to determine the basal heat flux immediately beneath the ice pack (*e.g.*, Harvey, 1990).

Figure 6 summarizes the algorithm to update ice cover, snow cover and water depth. The ice cover equations are applied when water temperature first drops below 0 °C. The ice thickness is set to its minimum value of 0.05 m, which is suggested by Patterson and Hamblin (1988) and Vavrus et al. (1996). The need for a minimum ice thickness relates primarily to horizontal variability of ice cover during the formation and closure periods. The ice cover equations are discontinued and open water conditions are restored in the model when the thermodynamic balance first produces ice thickness < 0.05 m.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

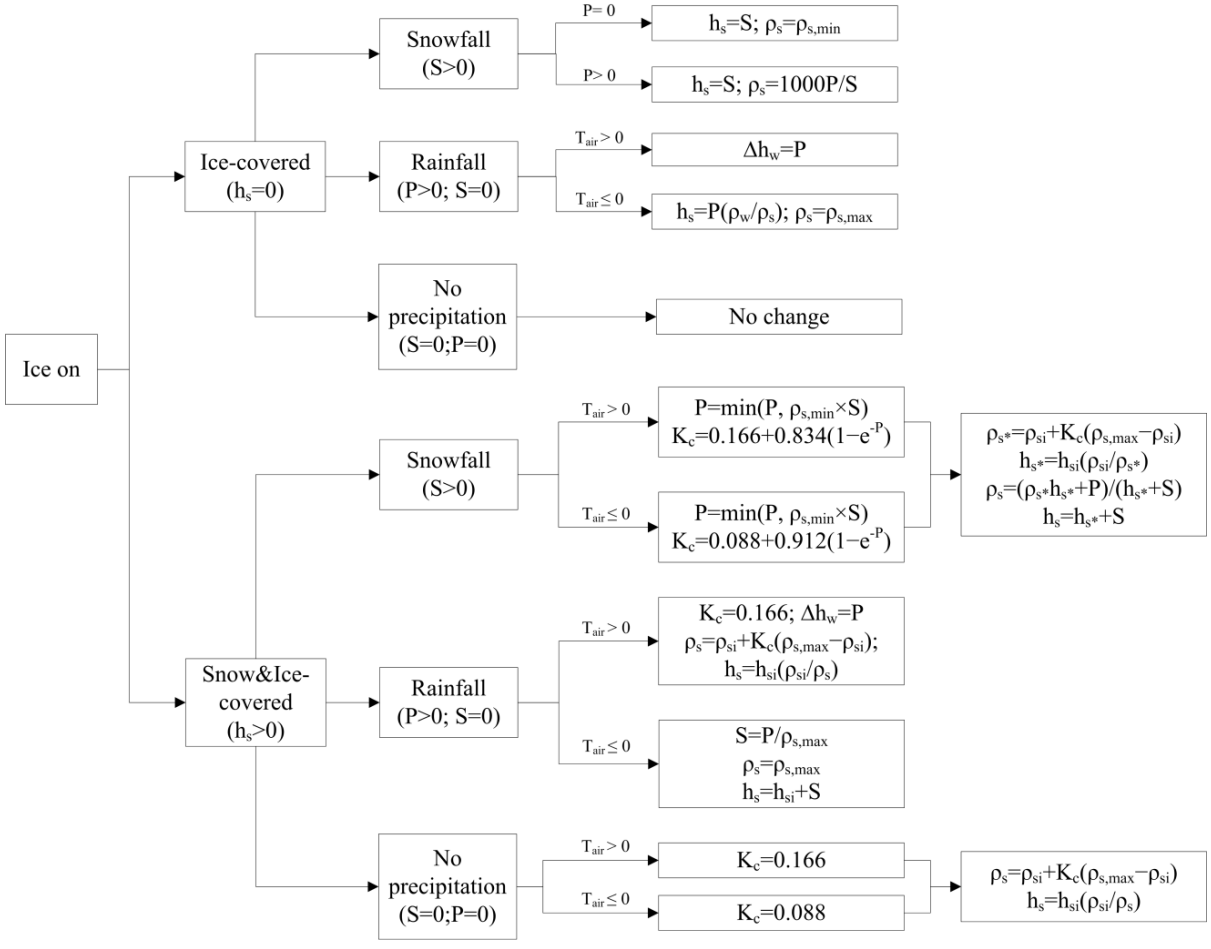

Figure 6: Decision tree to update ice cover, snow cover and water depth according to snow compaction, rainfall (*P*) and snowfall (*S*) on each day, and depth of snow cover (h_{si}) and snow density (ρ_{si}) for the previous day. Refer to Table 1 for definitions of other variables.

After the change in ice thickness due to heat exchange is calculated, the effects of snowfall, rainfall, and compaction of snow are calculated through appropriate choice of one of several options, depending on the air temperature and whether ice or snow is the upper solid boundary (Figure 6). Density of fresh snowfall is determined as the ratio of measured snowfall height to water-equivalent height, with any values exceeding the assigned maximum or minimum snow density (defaults: $\rho_{s,max}$ = 300 kg m^{-3}, $\rho_{s,min}$ = 50 kg m^{-3}) truncated to the appropriate limit. The snow compaction model is based on the exponential decay formula of McKay (1968), with selection of snow compaction parameters based on air temperature (Rogers et al., 1995) as well as on rainfall or snowfall. The approach of snow compaction used by Rogers et al. (1995) is to set the residual snow density to its maximum value when there is fresh snowfall. This method is found to produce increases in snow density that are too rapid when there is only light snowfall. As a result, GLM uses a gradual approach where the new snowfall and

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

the existing snow is used to form a layer with a combined mass and average density. Example outputs are shown in Figure 7, and see also Yao et al. (2014).

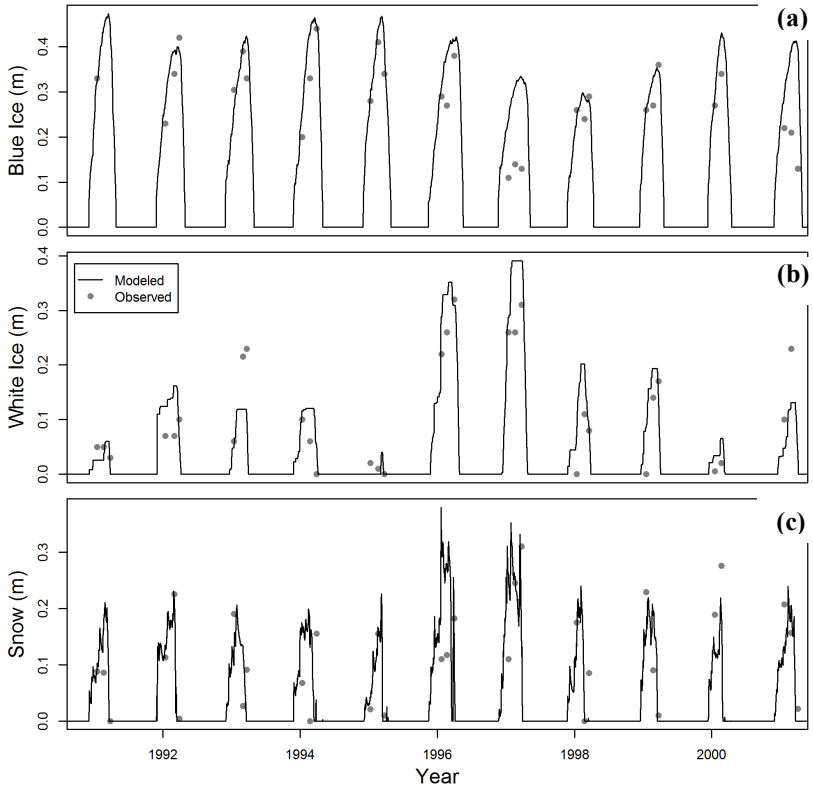

Figure 7: Example of modelled and observed blue ice (a), white ice (b) and snow (c) thickness for Sparkling Lake, Wisconsin. Lines are modelled thickness and points are average observed thicknesses.

2.5 Stratification and vertical mixing

2.5.1 Surface mixed layer

GLM works on the premise that the balance between the available energy, E_{TKE}, and the energy required for mixing to occur, E_{PE}, provides for the surface mixed layer (SML) deepening rate dz_{SML}/dt. For an overview of the dynamics readers are referred to early works on bulk mixed layer depth models by Kraus and Turner (1967) and Kim (1976), which were subsequently extended by Imberger and Patterson (1981) as a basis for hydrodynamic model design. In this model, the available kinetic energy is calculated due to contributions from wind stirring, shear production between layers, convective

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

overturn, and Kelvin-Helmholtz (K-H) billowing. They may be combined and summarised for E_{TKE} as (Hamilton and Schladow, 1997):

$$E_{TKE} = \underbrace{0.5C_K(w_*^3)\,\Delta t}_{convective\ overturn} + \underbrace{0.5C_K(C_W\,u_*^3)\,\Delta t}_{wind\ stirring} + \underbrace{0.5\,C_S\left[u_b^2 + \frac{u_b^2}{6}\frac{d\xi}{dz_{sml}} + \frac{u_b\xi}{3}\frac{du_b}{dz_{sml}}\right]}_{\substack{shear\ production \\ K-H\ production}}\Delta z_{k-1} \quad (30)$$

where ξ is the K-H billow length scale (described below), u_b is the shear velocity at the interface of the mixed layer, and C_K, C_W, and C_S are mixing efficiency constants. For mixing to occur, the energy must be sufficient to lift up water at the bottom of the mixed layer, denoted here as the layer $k-1$, with thickness Δh_{k-1}, and accelerate it to the mixed layer velocity. This also accounts for energy consumption associated with K-H production and expressed as, E_{PE}:

$$E_{PE} = \left[\underbrace{0.5C_T(w_*^3 + C_W\,u_*^3)^{2/3}}_{acceleration} + \underbrace{\frac{\Delta\rho}{\rho_o}\,g\,z_{SML}}_{lifting} + \underbrace{\frac{g\xi^2}{24\rho_o}\frac{d(\Delta\rho)}{dz_{sml}} + \frac{g\xi\Delta\rho}{12\rho_o}\frac{d\xi}{dz_{sml}}}_{K-H\ consumption}\right]\Delta z_{k-1} \quad (31)$$

where z_{SML} is the thickness of the surface mixed layer. To numerically resolve the above equations we sequentially compute the different components of the above expressions in light of the layer structure. Here GLM follows the algorithm outlined in Imberger and Patterson (1981) whereby cooling is computed so that layers are combined due to convection, then stirring, and then shear and K-H mixing are computed.

To compute mixing due to convective cooling, the value for w_* is calculated, which is the turbulent velocity scale associated with convection. The model adopts the algorithm used in Imberger and Patterson (1981), whereby the potential energy that is released by mixed layer deepening is computed by looking at the moments of the different layers in the surface mixed layer (from layers K to N_{LEV}):

$$w_*^3 = \frac{g}{\rho_{SML}\,\Delta t}\left(\sum_{k=K}^{N_{LEV}}\left[\rho_k\,\Delta z_k\,\widetilde{h_k}\right] - \widetilde{h_{SML}}\sum_{k=K}^{N_{LEV}}\left[\rho_k\,\Delta z_k\right]\right) \quad (32)$$

where ρ_{SML} is the mean density of the mixed layer including the combined layer, ρ_k is the density of the k^{th} layer, Δz_k is the height difference between two consecutive layers within the loop ($\Delta z_k = h_k - h_{k-1}$), $\widetilde{h_k}$ is the mean height of layers to be mixed ($\widetilde{h_k} = 0.5[\,h_k + h_{k-1}]$), and $\widetilde{h_{SML}}$ is the epilimnion (surface mixed layer) mid height, calculated from: $\widetilde{h_{SML}} = 0.5[\,h_{SURF} + h_{K-1}]$.

The velocity scale u_* is associated with wind stress and calculated according to the wind strength:

$$u_*^2 = C_D U_x^2 \quad (33)$$

where C_D is the drag coefficient for momentum. The model first has a check to see if the stirring energy can overcome the energy required to mix the k-1 layer, *i.e.*, mixing occurs if:

$$C_K(w_*^3 + C_W\,u_*^3)\,\Delta t \geq \left(g'_k\,z_{SML} + C_T(w_*^3 + C_W\,u_*^3)^{2/3}\right)\Delta z_{k-1} \quad (34)$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

and $g'_k = \frac{\Delta\rho}{\rho_o}$ is the reduced gravity between the mixed layer and k-1 layer. If the condition is not met the energy is stored for the next time-step.

Once stirring is completed, mixing due to velocity shear is applied. Velocity shear at the interface is approximated from:

$$u_b = \frac{{u_*}^2 t}{z_{sml}} + u_o \tag{35}$$

where t is characteristic time scale over which the shear has been operating, considered relative to t_{shear}, which is the time beyond which there is no shear production (*i.e.*, $u_b = 0$ if the time exceeds t_{shear}). This cut-off time assumes use of only the energy produced by shear at the interface during a period equivalent to half the basin-scale seiche duration, T_i, and modified to account for damping:

$$t_{shear} = T_i \left(1 + 0.59 \left[1 - cosh\left(\frac{T_d}{T_i} - 1\right)^{-1}\right]\right) \tag{36}$$

where T_d is the time-scale of damping (see Spigel, 1978). The wave period is approximated based on the stratification as $T_i = L_{META}/2c$, where L_{META} is the length of the basin at the thermocline and c is the internal wave speed. Once the velocity is computed, the energy for mixing from velocity shear is compared to that required for lifting and accelerating the next layer down, and layers are combined if there is sufficient energy:

$$0.5\, C_S \left[\frac{u_b^2(\widetilde{z_{SML}} + \Delta\xi)}{6} + \frac{u_b \xi \Delta u_b}{3}\right] + \left[g'_k \xi \left(\frac{\xi \Delta z_{k-1}}{24 z_{SML}} - \frac{\Delta\xi}{12}\right)\right] \geq \left(g'_k\, z_{SML} + C_T (w_*^3 + C_W\, u_*^3)^{2/3}\right) \Delta z_{k-1} \tag{37}$$

where the K-H length scale is $\xi = C_{KH} u_b^2 / g'_{EH}$ and $\Delta\xi = 2\, C_{KH}\, u_b\, \Delta u_b / g'_{EH}$; in this case the reduced gravity is computed from the difference between the epilimnion and hypolimnion, and C_{KH} is a measure of the billow mixing efficiency.

Once shear mixing is done, the model checks the resultant density interface to see if it is unstable to shear (i.e., K-H billows would be expected to form). This occurs if the gradient is less than the K-H length scale. If K-H mixing is required, layers are further split and subject to mixing using an algorithm similar to above.

2.5.2 Deep mixing

Mixing below the SML in lakes, in the deeper stratified regions of the water column, is modelled using a characteristic vertical diffusivity, $K_Z = K_\varepsilon + K_m$, where K_m is the fixed molecular diffusivity of scalars. Three hypolimnetic mixing options are possible in GLM including: (i) no diffusivity (ii) a constant vertical diffusivity K_Z over the water depth below the thermocline or (iii) a derivation by Weinstock (1981), which is described as being suitable for regions displaying weak or strong stratification, whereby diffusivity increases with dissipation and decreases with heightened stratification.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

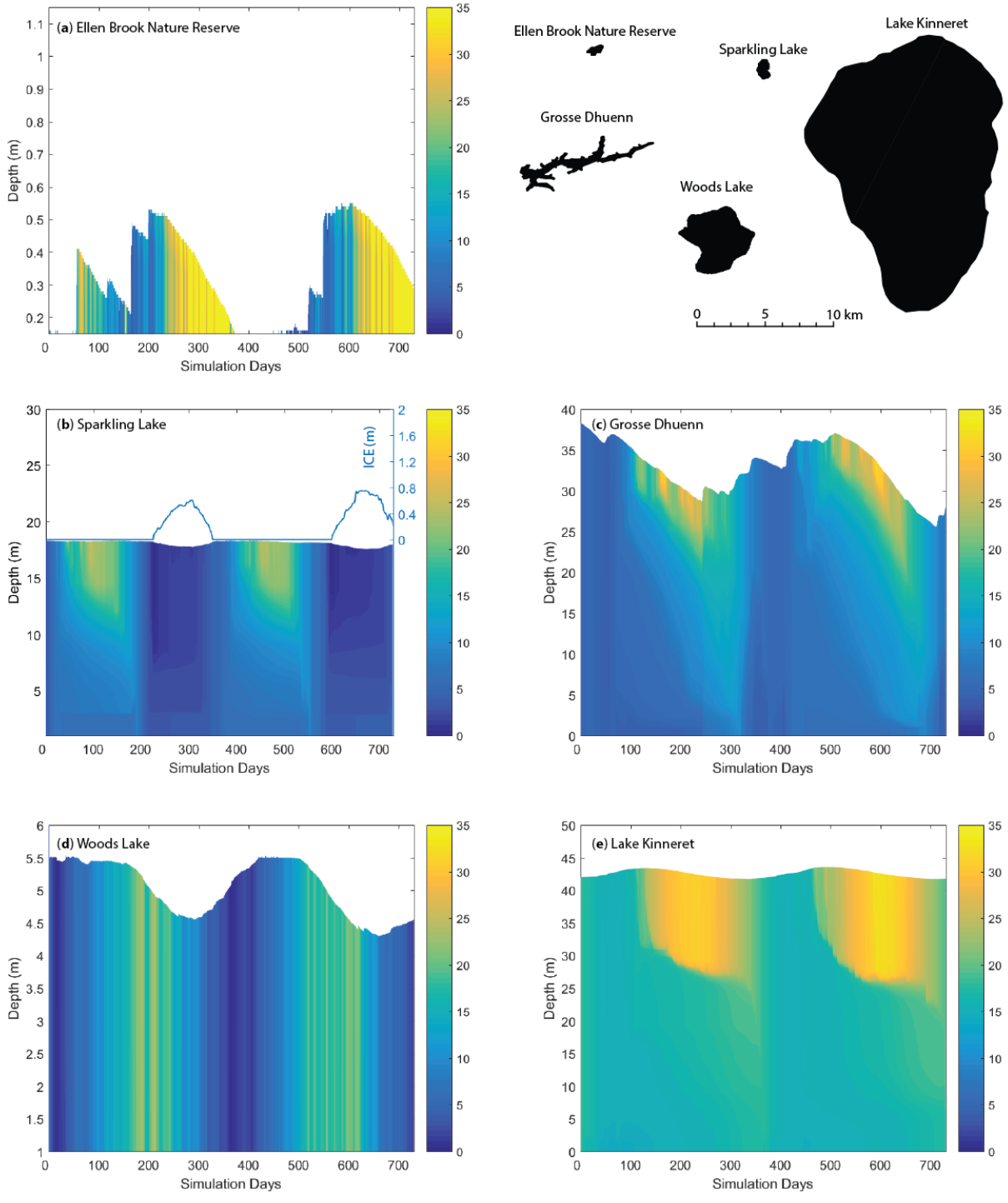

Figure 8: A two-year times-series of the simulated temperature profiles for five example lakes, a-e, that range in size and hydrology. For more information about each lake and the simulation configuration refer to the Data availability section (refer also to Fig. 2 and 5). Sparkling Lake (d) also indicates the simulated depth of ice.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

For the constant vertical diffusivity option, the coefficient α_{TKE} is interpreted as a vertical diffusivity (m²/s). For the Weinstock (1981) model, the diffusivity is computed according to:

$$K_z = \frac{\alpha_{TKE}\varepsilon_{TKE}}{N^2 + 0.6\,{k_{TKE}}^2 {u_*}^2} \tag{38}$$

where α_{TKE} is the mixing efficiency of hypolimnetic TKE (~0.8 in Weinstock, 1981) and k_{TKE} is the turbulence wavenumber:

$$k_{TKE} = \frac{12.4\,A_{top}}{\tilde{V}\,\Delta z_{top}\,10^3} \tag{39}$$

and $u_* = \sqrt{1.612 \times 10^{-6}\,{U_x}^2}$. The term N^2 is the Brunt–Väisälä (buoyancy) frequency defined as:

$$N^2 = \frac{g\Delta\rho}{\rho\Delta z} \approx \left[\frac{g(\rho_{i+2} - \rho_{i-2})}{\rho_{ref}(h_{i+2} - h_{i-2})}\right] \tag{40}$$

Estimating the turbulent dissipation rate can be complex and GLM adopts a simple approach as described in Fischer et al. (1980) where a "net dissipation" is approximated by assuming dissipation is in equilibrium with energy inputs from external drivers:

$$\varepsilon_{TKE} \approx \overline{\varepsilon_{TKE}} = E_{WIND} + E_{INFLOW} \tag{41}$$

which is expanded and calculated per unit volume as:

$$\varepsilon_{TKE} = \underbrace{\frac{1}{(\tilde{V}\bar{\rho})10^3}\frac{m\,C_D\rho_a\,f_S\,{U_x}^3\,A_l}{10^6}}_{\text{rate of working by wind}} + \underbrace{\frac{1}{(V_{mix}\bar{\rho})10^3}\sum_{i}^{N_{INF}} g\,\Delta\rho_i\,Q_i(h_{top} - h_i)}_{\text{rate of work done by inflows}} \tag{42}$$

The diffusivity is calculated according to Eq. 42, but since the dissipation is assumed to concentrate close to the level of strongest stratification, the "mean" diffusivity is modified to decay exponentially with distance from the thermocline:

$$K_{z_l} = \begin{cases} 0 & h_l \geq (h_{top} - z_{mix}) \\ K_z\ exp\left[\dfrac{-(h_{top} - z_{mix} - h_l)^2}{\sigma}\right] & h_i < (h_{top} - z_{mix}) \end{cases} \tag{43}$$

where σ is the variance the N^2 distribution below h_{mix} and scales with the depth over which mixing decays.

Once the diffusivity is approximated (for either model 1 or 2 in Eq. 43), the diffusion of any scalar, C, between two layers is numerically accounted for by the following mass transfer expressions:

$$C_{i+1} = \bar{C} + \frac{\exp(-f)\Delta z_i \Delta C}{(\Delta z_{i+1} + \Delta z_i)}$$
$$C_i = \bar{C} - \frac{\exp(-f)\Delta z_{i+1} \Delta C}{(\Delta z_{i+1} + \Delta z_i)} \tag{44}$$

where $\bar{C}$ is the weighted mean concentration of C for the two layers, and ΔC is the concentration difference between them. f is related to the diffusivity according to:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

$$f = \frac{K_{z_{i+1}} + K_{z_i}}{(\Delta z_{i+1} + \Delta z_i)^2} \Delta t \tag{45}$$

The above diffusion algorithm is run once up the water column and once down the water column as a simple explicit method for capturing diffusion to both the upper and lower layers.

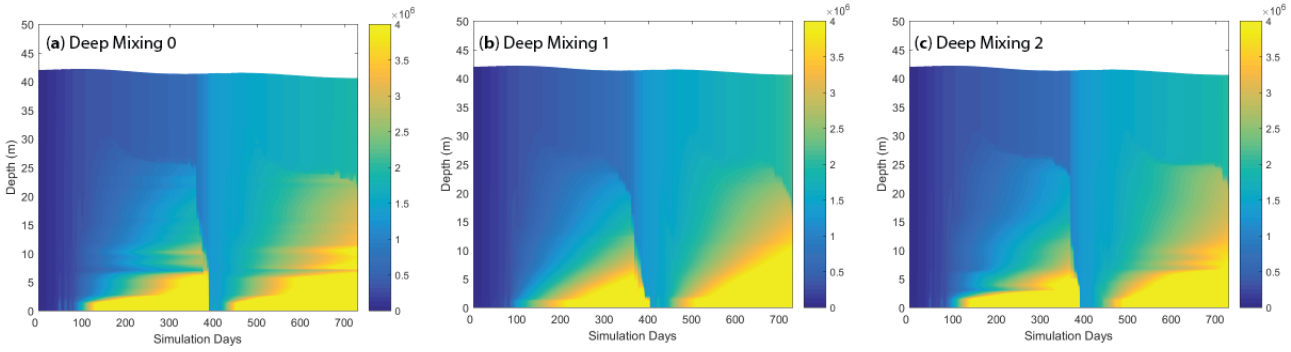

Figure 9: Simulations for Lake Kinneret showing hypolimnetic concentration of a passive tracer released from the bottom sediment at a constant rate for the case a) without deep mixing, b) constant vertical diffusivity, and c) calculated vertical diffusivity (Eq. 38). For thermal structure of this case refer to Figure 8c.

2.6 Inflows and outflows

Inflows can be specified as local runoff from the surrounding (dry) lake domain (Q_R described above, Eq. 5), rivers entering at the surface of the lake that will be buoyant or plunge depending on their momentum and density (Sect 2.6.1), or submerged inflows including groundwater (Sect 2.6.2). Any number of inflows to the lake body can be specified and these are applied daily. Four options for outflows are included in GLM, including withdrawals from a specified depth (Sect 2.6.3), adaptive offtake (Sect 2.6.4), vertical groundwater seepage (Sect 2.6.5).

2.6.1 River inflows

For river inflows, depending on the density of the river water, the inflow will form a positively or negatively buoyant intrusion that will enter the lake and insert at a depth of neutral buoyancy. As the inflow inserts it will entrains water depending on the rate of mixing created by th inflowing water. In GLM, as the inflow crosses layers it will entrain water from each, until it reaches a level of neutral buoyancy and undergoes insertion. Therefore, when it reaches its point of neutral buoyancy a new layer of thickness dependent on the inflow volume at that time (including additions from entrainment) is created. Following insertion, the inflow layer may then amalgamate with adjacent layers depending on numerical criteria within the model for combining or splitting layers.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

The rate of entrainment of the intrusion, E, can be calculated in a number of ways. For simplicity, in GLM the rate has been adapted from the first approximation given in Fischer et al. (1979):

$$E = 1.6\frac{{C_{D_i}}^{3/2}}{Ri} \tag{46}$$

where C_{D_i} is the user specified drag coefficient for the inflow. The Richardson's number is adapted from Fischer et al. (1979) as:

$$Ri = \frac{C_{D_i}\left(1 + 0.21\sqrt{C_{D_i}}\,\sin\alpha_{inf}\right)}{\sin\alpha_{inf}\tan\phi_{inf}} \tag{47}$$

where α_{inf} is the stream half angle and ϕ_{inf} is the slope of the inflow at the point where it meets the water body (Figure 10).

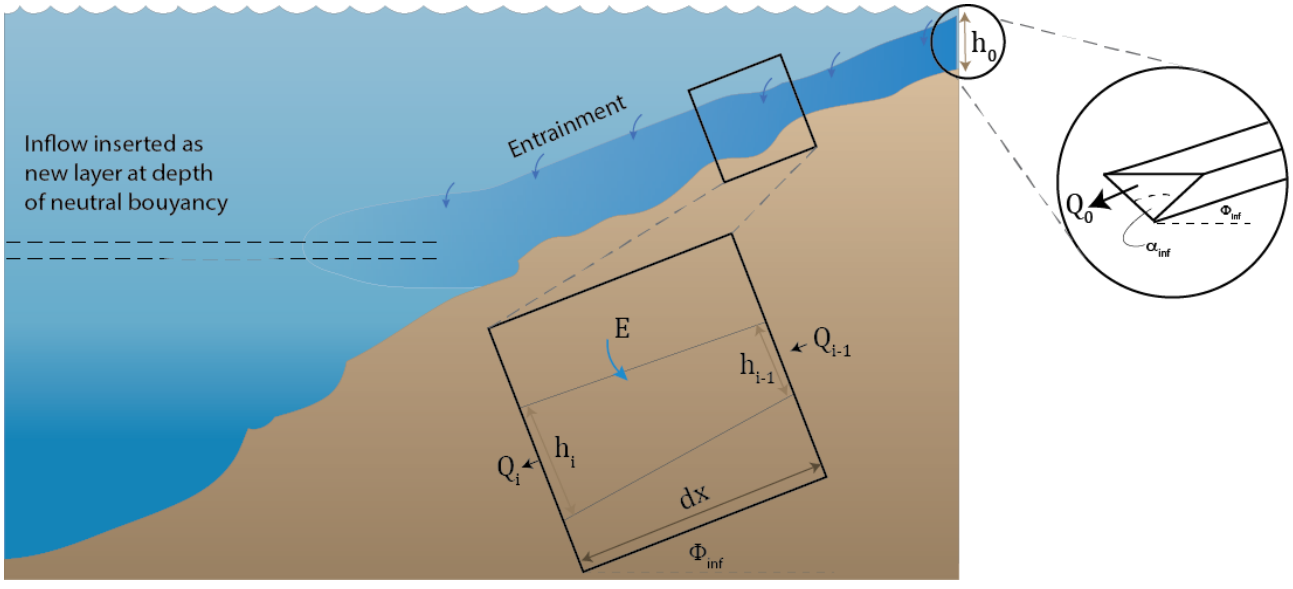

Figure 10: Schematic showing inflow insertion height, entrainment, E, slope, ϕ_{inf} and half angle, α_{inf} of an inflowing river entering with a user prescribed flow of Q_0, and estimated starting height of h_0.

As the inflow parcel travels through the layers, the increase in inflow thickness due to entrainment is estimated as:

$$h_i = 1.2Edx + h_{i-1} \tag{48}$$

where h_i is the inflow thickness, E is the entrainment rate and dx is the distance travelled by the inflowing water, calculated from the flow rate and inflow thickness. The initial estimate of the intrusion height is computed from Imberger and Patterson (1981) and Antenucci et al. (2005):

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

$$h_0 = \left(2{Q_{inf}}^2 \frac{Ri}{{g'}_{inf}} tan^2\phi_{inf}\right)^{1/5} \quad (49)$$

where Q_{inf} is the inflow discharge provided as a boundary condition and g' is the reduced gravity of the inflow given as:

$$g'_{inf} = g\frac{(\rho_{inf} - \rho_s)}{\rho_s} \quad (50)$$

where ρ_{inf} is the density of the inflow and ρ_s the density of the surface layer. The distance travelled by the inflow aliquot, dx, is estimated as the distance travelled in the vertical and the slope of the inflow river, ϕ_{inf} and given by:

$$dx = \frac{dz}{\sin\phi_{inf}} \quad (51)$$

where dz is the distance travelled in the vertical. The velocity of the inflow aliquot for that day is then calculated as:

$$u = h_i^2 \frac{Q_{inf}}{\tan\alpha} \quad (52)$$

Following conservation of mass, the flow is estimated to increase according to Fischer et al. (1979), as in Antenucci et al. (2005):

$$Q_i = Q_{i-1}\left[\left(\frac{h_i}{h_{i-1}}\right)^{5/3} - 1\right] \quad (53)$$

The above entrainment and insertion algorithm is repeated for each inflow. Aside from importing mass into the lake, river inflows also contribute turbulent kinetic energy to the hypolimnion, as discussed in the Sect 2.5.2 (e.g., see Eq. 42), and contribute to the scalar transport in the water column (Figure 11a).

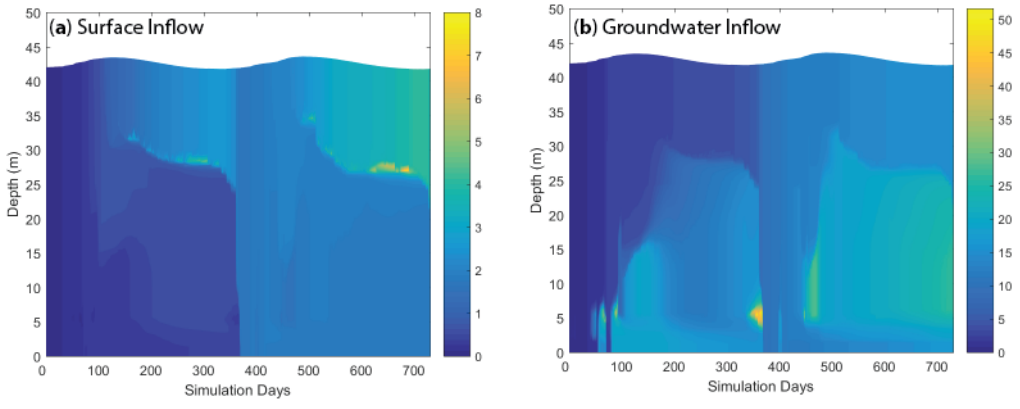

Figure 11: Simulation showing inflow tracer insertion example for the case where a) the inflow was set as a surface river inflow, and b) the inflow was set as a submerged inflow at a specified height (h=5m). After input the tracer is subject to mixing during inflow entrainment and by surface and deep mixing once inserted.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

2.6.2 Submerged inflows

Submerged inflows are inserted at the user-specified depth with zero entrainment. The submerged inflow volume is added as a new layer which may then be mixed with adjacent layers (above or below) depending on the density difference, until neutral buoyancy is attained (Figure 11b). This option can be used across one or more layers to account for groundwater inputs, or for capturing a piped inflow, for example.

2.6.3 Withdrawals

Outflows from a specific depth can be accommodated including outlets from a dam wall offtake, or other piped withdrawal, or for removing water that may be lost due to groundwater discharge. For a stratified water column, the water will be removed from the layer corresponding to the specified withdrawal depth, as well as layers above or below depending on the strength of discharge and stability of the water column. Accordingly, the model assumes a line-sink algorithm where the thickness of the withdrawal layer is dependent on the internal Froude (Fr) and Grashof (Gr) numbers, and the parameter, R (see Fischer et al., 1979; Imberger and Patterson, 1981):

$$Fr = \frac{Q_{outf}}{N_{outf}^2 W_{outf} L_{outf}^2} \tag{54}$$

$$Gr = \frac{N_{outf}^2 A_{outf}^2}{v_{outf}^2} \tag{55}$$

$$R = FrGr^{1/3} \tag{56}$$

where W_{outf}, L_{outf} and A_{outf} are the width, length and area of the lake at the outlet elevation, and v_{outf}^2 is the vertical diffusion coefficient averaged over the withdrawal layer. The Brunt- Väisälä frequency averaged over the thickness of the withdrawal layer, N_{outf}^2, is calculated as:

$$N_{outf}^2 = \frac{g}{dz}\frac{\rho_{outf} - \rho_i}{\rho_{outf}} \tag{57}$$

where dz is the thickness of the withdrawal layer, ρ_{outf} is the density of the lake at the height of withdrawal and ρ_i is the density of the lake at the edge of the withdrawal layer. The thickness of the withdrawal layer is then calculated depending on the value of R (Fischer et al. 1978), such that:

$$\delta_{outf} = \begin{cases} 2L_{outf} Gr^{-1/6} & R \leq 1 \\ 2L_{outf} Fr^{1/2} & R > 1 \end{cases} \tag{58}$$

The proportion of fluid withdrawn from each layer either above or below the layer of the outlet elevation is determined using a curve that fits the region of fluid drawn in a given time. To calculate the width and length of the lake at the height of outflow it is assumed, firstly, that the lake shape can be approximated as an ellipse, and secondly, that the ratio of length to width at the height of the outflow is the same as that at the lake crest. The length of the lake at the outflow height, L_{outf} and the lake width, W_{outf} are given by:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

$$L_{outf} = \sqrt{A_{outf} \frac{4}{\pi} \frac{L_{crest}}{W_{crest}}} \tag{59a}$$

$$W_{outf} = L_{outf} \frac{W_{crest}}{L_{crest}} \tag{59b}$$

where A_{outf} is the area of the lake at the outflow height, L_{crest} is the length and W_{crest} the width of the lake at the crest height. Depending on the layer(s) the water is withdrawn from, the water taken will have the associated scalar concentrations.

2.6.4 Adaptive offtake dynamics

For reservoir applications, a special outflow option has been implemented that extends the dynamics in Sect. 2.6.3 to simulate an adaptive offtake or selective withdrawal. This approach is used for accommodating flexible reservoir withdrawal regimes and their effects on the thermal structure within a reservoir. For this option, a target temperature is specified by the user and GLM estimates the corresponding withdrawal height within a predefined (facility) range to meet this target temperature during the runtime of the simulation, i.e., the withdrawal height adaptively follows the thermal stratification in the reservoir. The target temperature can be defined as a constant temperature (e.g., 14 °C) or a time-series such as a measured water temperature from an upstream river (via a *.csv file). The height of the adaptive offtake is printed out in a *.txt file and may be used for reservoir operation. In addition to the basic adaptive offtake function, GLM can also simulate withdrawal mixing, i.e., water from the adaptive offtake is mixed with water from another predefined height (e.g., the bottom outlet). For this option, the discharges at both locations need to be predefined by the user (via outflow *.csv file) and GLM chooses the adaptive withdrawal from a height, where the water temperature is such that the resulting mixing temperature meets the target temperature. This withdrawal mixing is a common strategy in reservoir operation where deep water withdrawal and temperature control are required simultaneously.

An example of the adaptive offtake function with and without withdrawal mixing, assuming a constant water temperature of 14 °C for the outflow water, shows that GLM is able to deliver a constant outflow temperature of 14 °C during the stratified period (Figure 12). In winter, when the water column is cooler than 14 °C, the model withdraws surface water. The adaptive offtake functionality can be used in a stand-alone mode or coupled to the dissolved oxygen concentration (via the water quality model AED2). In the latter case, the effect of the withdrawal regime on the oxygen dynamics in the hypolimnion can be simulated (see Weber et al., 2017). In this setting, the simulated hypolimnetic dissolved oxygen concentration at a specified height is checked against a critical threshold. If the hypolimnetic oxygen falls below the critical threshold, the height of the adaptive offtake will be automatically switched to a defined height (usually deep outlets in order to get rid of the oxygen-depleted water) to withdraw water from this layer, until the oxygen concentrations have recovered.

(a) **(b)**

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

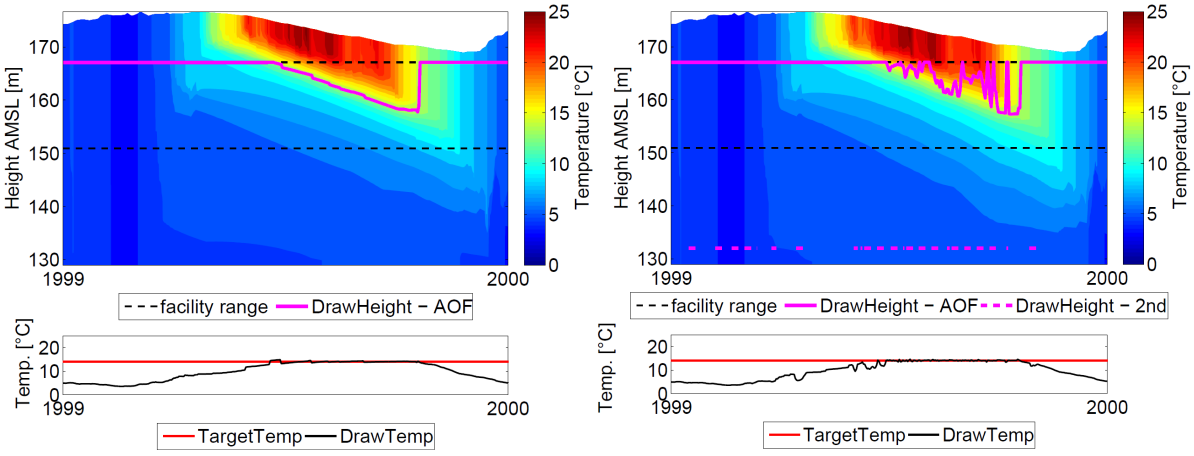

Figure 12: Adaptive offtake reservoir simulation; water temperatures of the adaptive offtake model assuming a constant temperature of 14 °C without (a) and with (b) mixing with bottom outlet withdrawal. The black dashed line represents the range of the variable withdrawal facility and the magenta lines the adaptive offtake and second withdrawal height.

2.6.5 Seepage

Seepage of water from the bottom layer is also configurable within the model, for example, as might be required in a wetland simulation or for small reservoirs perched above the water table that experience leakage to the soil below. Seepage is configured to leave the lake at a constant rate:

$$\frac{dh_B}{dt} = -G \tag{60}$$

where h_B is the depth of the bottom-most layer at any time, and G is the seepage rate (m day^{-1}). G is constrained within the model to ensure no more than 50% of the layer can be reduced in any one time-step. Note that in shallow-lake or wetland simulations, the layer structure may simplify to a single layer, in which case the surface and bottom layer are the same, and Eqs. 4 and 60 are effectively combined.

2.7 Wave height and bottom stress

Wind-induced resuspension of sediment from the bed of shallow lakes is sporadic and occurs as the waves created at the water surface create oscillatory currents that propagate down to the lake-bed. GLM does not predict resuspension and sediment concentration directly, but computes the bottom shear stress for later use by sediment and water quality modules. Nonetheless, even without this explicit formulation, the model can identify the areal extent and potential for bed-sediment resuspension by computing the area of the lake over which the bed shear stress exceeds some critical value required for resuspension to occur.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

To compute the stress at the lake bottom we estimate the surface wave conditions using a simple, fetch-based, steady state wave model (Laenen and LeTourneau, 1996; Ji 2008). The wave geometry (wave period, significant wave height and wave length), is predicted based on the wind speed and fetch over which the waves develop (Figure 13), calculated as:

$$F = 2\sqrt{A_s/\pi} \tag{61}$$

Using this model, the wave period, T, is calculated from fetch as:

$$T = 7.54\left(\frac{U_x}{g}\right) tanh(\xi)\, tanh\left(\frac{0.0379\left[\frac{gF}{U_x^2}\right]^{0.333}}{tanh(\xi)}\right) \tag{62}$$

where:

$$\xi = 0.833\left[\frac{gd_{avg}}{U_x^2}\right]^{0.375} \tag{63}$$

and h_{avg} is the average lake depth. Wave length is then estimated from:

$$L = \left[\frac{gT^2}{2\pi}\right] tanh\left(\frac{2\pi\, d_{avg}}{\left[\frac{gT^2}{2\pi}\right]}\right) \tag{64}$$

and wave height from:

$$H_s = 0.283\left(\frac{U_x^2}{g}\right) tanh(\zeta) tanh\left(\frac{0.00565\left[\frac{gF}{U_x^2}\right]^{0.5}}{tanh(\zeta)}\right) \tag{65}$$

where

$$\zeta = 0.53\left[\frac{gd_{avg}}{U_x^2}\right]^{0.75} \tag{66}$$

Based on these properties the orbital wave velocity at depth (in the i^{th} layer) is calculated as:

$$U_{orb_i} = \frac{\pi H_s}{T sinh\left[\frac{2\pi\, d_i}{L}\right]} \tag{67}$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

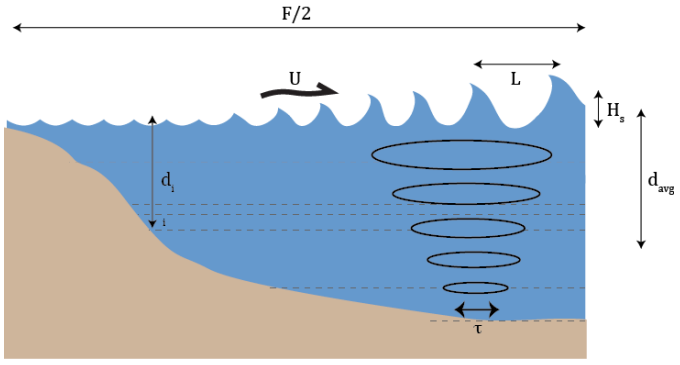

Figure 13: Schematic of the wave estimation approach depicting the lake fetch, surface wind speed, wave height and wavelength, and bottom stress created by the orbital velocity.

The total shear stress at the lake bed is calculated as:

$$\tau_i = \frac{1}{2}\,\rho_w\left[\,f_w\,{U_{orb_i}}^2 + \,f_c\,{U_{m_i}}^2\right] \tag{68}$$

where U_m is the mean velocity of the layer, computed during the mixing calculations (Eq. 33). The friction factors use D (a typical particle diameter). For the current stress we compute $f_w = \,0.24/\log\left(12d_{avg}/2.5D\right)$ and for waves:

$$f_w = \exp\left[-5.977 + 5.213\left(\frac{a}{2.5D}\right)^{-0.19}\right] \quad \text{Option 1 : Laenen and LeTourneau, 1996} \tag{69}$$

$$f_w = 0.00251\exp\left[5.213\left(\frac{U_{orb}\,T}{4\pi D}\right)^{-0.19}\right] \quad \text{Option 3 : Kleinhans \& Grasmeijer (2006)}$$

$$f_w = \frac{2\beta g\rho_D D}{U^2\rho_w} \quad \text{Option 3 : Le Roux (2007)}$$

Figure 14 demonstrates wave-related outputs from a shallow lake.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

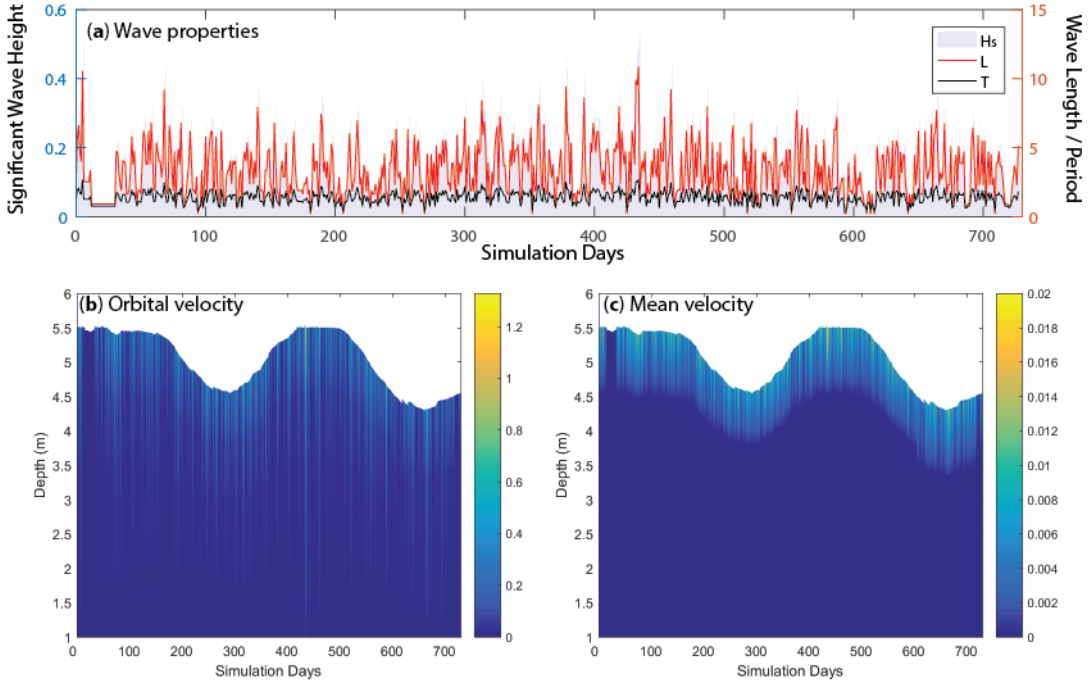

Figure 14: Simulation from Woods Lake, Australia, showing a) time-series of surface wave properties (H_s, L and T), b) orbital velocity, U_{orb}, changes over time (m/s), and c) comparison with the mean velocity, U_m (m/s).

3 Code organization and model operation

Aside from the core water balance and mixing functionality, the model features numerous options and extensions in order to make it a fast and easy-to-use package suitable for a wide range of contemporary applications. Accommodating these requirements has led to the code structure outlined in Figure 15. The model is written in C, with a Fortran-based interface module to link with Fortran-based water quality modelling libraries in Sect. 4. The model compiles with gcc, and gfortran, and commercial compilers, with support for Windows, OS X and Linux.

To facilitate the use of the model in teaching environments and for users with limited technical support, the model may be operated without any third party software, as the input files consist of "namelist" (`nml`) text files for configuration and csv files for meteorological and flow time-series data (Figure 16). The outputs from predictions are stored into a structured netcdf file, and this can be visualised in real-time through the simple inbuilt plotting library (libplot), or may be opened for post-processing in MATLAB or R (see Sect. 5.1). Parameters and configuration details are input through the main `glm.nml` text file (Figure 16) and default parameters and their associated description are outlined in Table 1.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

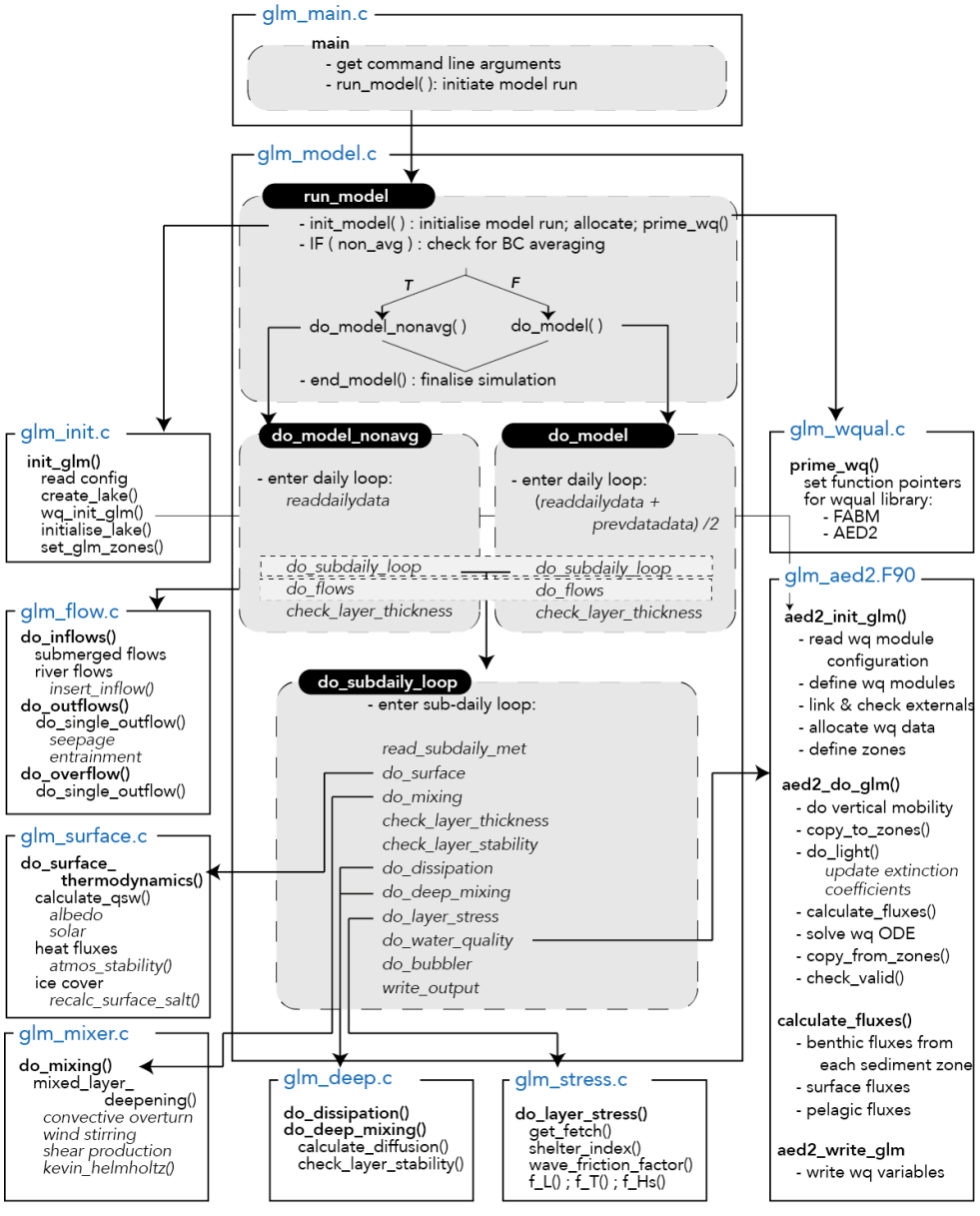

Figure 15: GLM code structure and logic flow. Each module is depicted as a box with the main routines and functions summarised.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

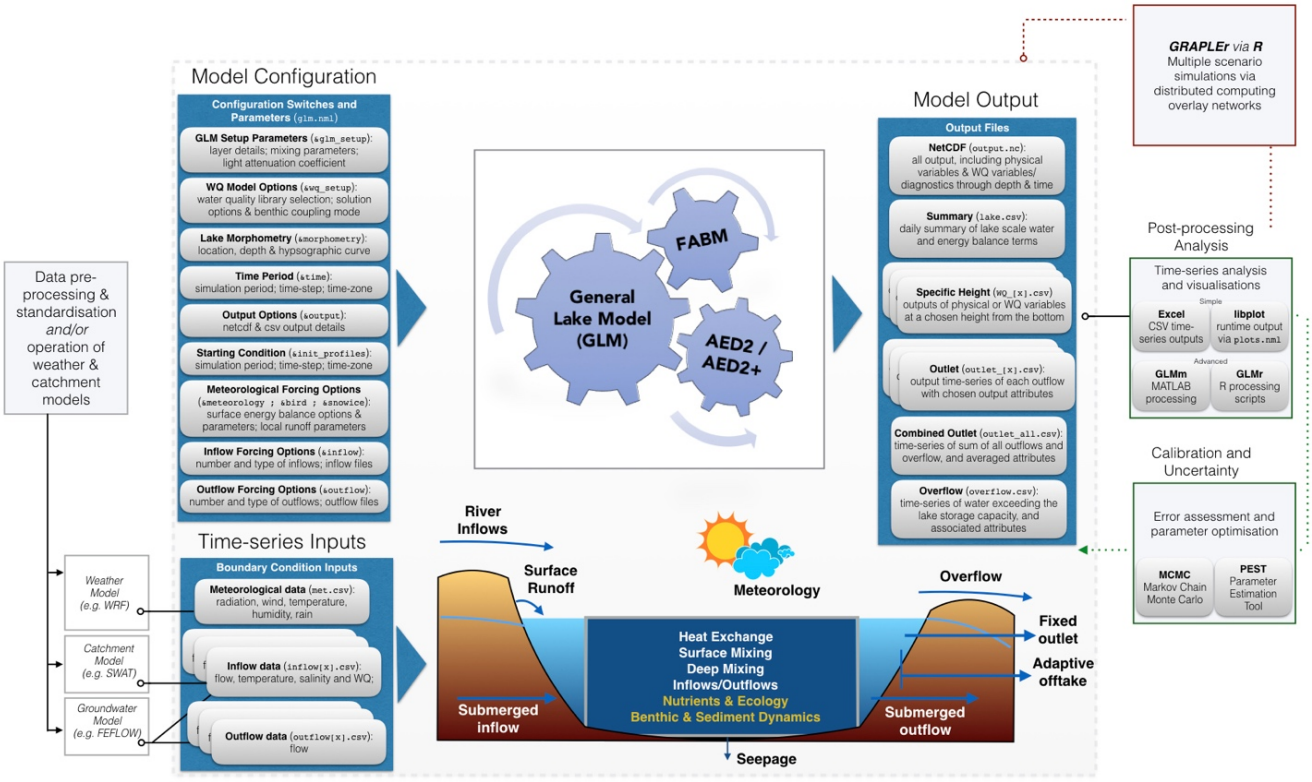

Figure 16: Flow diagram showing the input information required for operation of the model, and outputs and analysis pathways.

4 Dynamic coupling with biogeochemical and ecological model libraries

Beyond modelling the vertical temperature distribution, the water, ice and heat balance, as well as the transport and mixing in a lake, the model has been designed to couple with biogeochemical and ecological model libraries. Currently the model is distributed pre-linked with the AED2 simulation library (Hipsey et al., 2013) and the Framework for Aquatic Biogeochemical Models (FABM; Bruggeman and Bolding, 2014). Through connection with these libraries, GLM can simulate the seasonal changes in vertical profiles of turbidity, oxygen, nutrients, phytoplankton, zooplankton, pathogens and other water quality variables of interest. Documentation of these models is beyond the scope of the present paper, however, two features are highlighted here relevant to managing physical-ecological interactions.

Firstly, the model is designed to allow a user-defined number of sediment zones that span the depth of the lake. Using this approach, the current setup allows for depth-dependent sediment properties, both for physical properties such as roughness or sediment heat flux, and also biogeochemical properties such as sediment nutrient fluxes and benthic ecological

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

interactions. Since the GLM layer structure is flexible over time (i.e., layer depths are not fixed), any interactions between the water and sediment/benthos must be managed at each time-step. The model therefore supports disaggregation and/or aggregation of layer properties, for mapping individual water layers to one or more sediment zones (Figure 17). The weightings provided by each layer to the sediment are based on the relative depth overlap of a layer with the depth range of the sediment zone. This approach makes the model suitable for long-term assessments of wetland, lake and reservoir biogeochemical budgets, including for C, N and other attribute balances as required (Stepanenko et al., 2016).

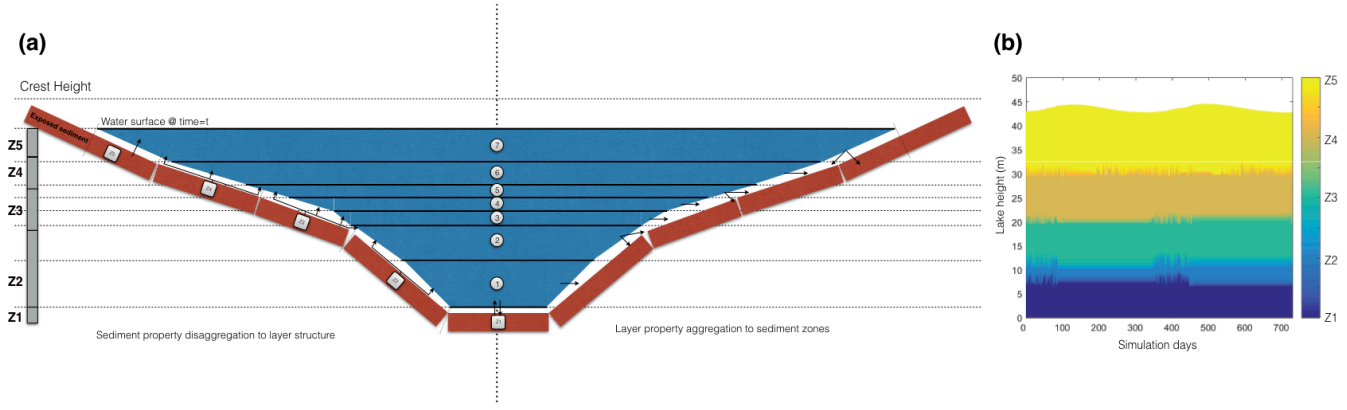

Figure 17: a) Schematic of a lake model layer structure (indicated by layers 1-7), in conjunction with five sediment "zones" (Z1-Z5) activated when `benthic_mode` = 2. The dynamically varying layer structure is re-mapped to the fixed sediment zone locations at each time step in order for the sediment zone to receive the average overlying water properties, and for the water to receive the appropriate information from benthic/sediment variables. b) example of GLM output showing the sediment zone each water layer is mapped to.

Secondly, the water quality modules feed back to GLM properties related to the water and/or heat balance. Feedback options include water density additions, bottom drag, f_w, the light attenuation coefficient, K_w, solar shading and rainfall interception.

5 Workflow tools for integrating GLM with sensor data and supporting models

The GLM model has been designed to support integration of large volumes of data coming from instrumented lakes, including many GLEON sites (cite ref, e.g., Weathers, Hanson, etc.). These data consist of high-frequency and discrete time-series observations of hydrologic fluxes, meteorology, temperature, and water quality (e.g., Hamilton et al. 2014). To facilitate research that requires running the model using these data sources, we have created GLM interfaces in the R and MATLAB analysis environments. These tools support user-friendly access to the model and include routines that streamline the process of calibrating models or running various scenarios. In addition, for assessment of lake dynamics in response to catchment or climatic forcing it is desirable to be able to connect GLM with other model platforms associated with surface and groundwater simulation, and weather prediction (Read et al., 2016).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

5.1 R and MATLAB libraries for model setup and post-processing

The R and MATLAB scientific languages are commonly used in aquatic research, often as part of automated modelling and analysis workflows. GLM has a client library for both, and these tools are shared freely online. The R package is called "`glmtools`" (https://github.com/GLEON/glmtools) and the MATLAB library is called "`GLMm`" (https://github.com/AquaticEcoDynamics/GLMm). Both tools have utilities for model pre- and post-processing. The pre-processing components can be used to format and modify data inputs and configuration files, and define options for how GLM executes. Post-processing tools include visualizations of simulation results (as shown in the results figures above), comparisons to field observations, and various evaluations of model performance.

5.2 Utilities for assessing model performance, parameter identification and uncertainty analysis

In order to compare the performance of the model for varied types of lakes, numerous different metrics of model performance are relevant. These include simple measures like surface or bottom temperature, or ice thickness, however, it is also possible to compare the model's performance in capturing higher-order metrics relevant to lake dynamics, including Schmidt Stability, thermocline depth, ice on/off dates (see also Bruce et al., 2017, for a detailed assessment of the model's accuracy across a wide diversity of lakes across the globe). With particular interest in the model's ability to interface with high frequency sensor data for calculation of key lake stability metrics (Read et al., 2011), then continuous wavelet transform comparisons are also possible (Kara et al., 2012), allowing assessment of the time-scales over which the model is able to capture the observed variability within the data.

As part of the modelling process, it is common to adjust parameters to get the best fit with available field data and, as such, the use of a Bayesian Hierarchical Framework in the aquatic ecosystem modelling community has become increasingly useful (e.g., Zhang and Arhonditsis, 2009; Romarheim et al., 2015). Many parameters described throughout Sect. 2 are attempts at physically based descriptions where there is relatively little variation (Bruce et al., 2017), thereby reducing the number of parameters that remain uncertain, however, for others their variation reflects imperfect formulation of some processes that are not fully considered. Therefore, within MATLAB, support scripts for GLM to work with the Markov Chain Monte Carlo (MCMC) code outlined in Haario et al. (2006) can be used to provide improved parameter estimates and uncertainty assessment (Figure 18). Wrappers and examples for use of GLM within the openDA framework and PEST are also being tested, giving users access to a wide range of uncertainty assessment and data assimilation algorithms. The PEST framework allows for calibration of complex model using highly-parameterised regularisation with pilot-points (Doherty, 2015), and sensitivity matrices derived from the calibration process can also be utilised in linear and non-linear uncertainty analysis.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

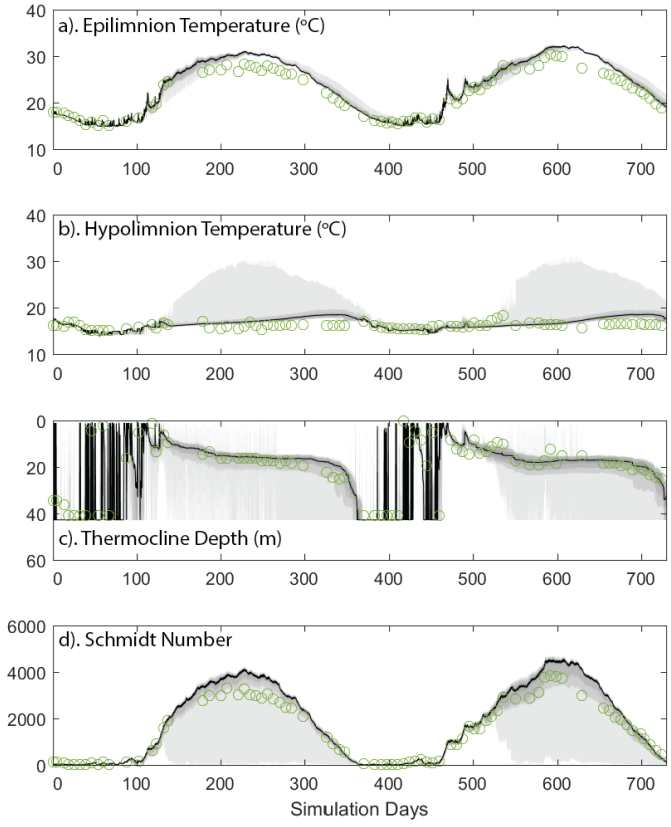

Figure 18: Depiction of parameter uncertainty for a GLM simulation of Lake Kinneret, Israel, following calibration against observations (green circles) via MCMC for a) epilimnion temperature, b) hypolimnion temperature, c) thermocline depth, and d) Schmidt number. The black line indicates the mean likelihood of the prediction, and the grey bands depict the 40th, 60th and 80th percentile.

5.3 Operation in the cloud: GRAPLEr

Questions relevant to land use and climate change are driving scientists to develop scenarios for how lake ecosystem services might respond to changing exogenous drivers. An important approach to addressing these questions is to simulate lake or reservoir physical-biological interactions in response to changing hydrology, nutrient loads or meteorology, and then infer consequences from the emergent properties of the simulation, such as changes in water clarity, extent of anoxia, mixing regime, or habitability to fishes (Hipsey et al., 2015). Often, it takes years or even decades for lakes to respond fully to changes in exogenous drivers, requiring simulations to recreate lake behavior over extended periods. While most desktop computers can run a decade-long, low-resolution simulation in less than one minute, high-resolution simulations of the same

extent may require minutes to hours of processor time. When questions demand hundreds, thousands or even millions of simulations, the desktop approach is no longer suitable.

Through access to distributed computing resources, modellers can run thousands of GLM simulations in the time it takes to run a few simulations on a desktop computer. Collaborations between computer scientists in the Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) and GLEON have led to the development of GRAPLEr (GLEON Research and PRAGMA Lake Expedition in R), software, written in R, that enables modellers to distribute batches of GLM simulations to pools of computers (Subratie et al., 2017). Modellers use GRAPLEr in two ways: by submitting a single simulation to the GRAPLEr Web service, along with instructions for running that simulation under different climate scenarios, or by configuring many simulations on the user's desktop computer, and then submitting them as a batch to the Web service. The first approach provides a high degree of automation that is well suited to training and instruction, and the second approach has the full flexibility often needed for research projects. In all approaches, GRAPLEr converts the submitted job to a script that is used by HTCondor (Thain et al., 2005) to distribute and manage jobs among the computer pool and ensure that all simulations run and return results. An iPOP overlay network (Ganguly et al., 2006) allows the compute services to include resources from multiple institutions, as well as cloud computing services. GRAPLEr's Web service front-end shields the modeller from the compute environment, greatly reducing the need for modellers to understand distributed computing; they therefore only need to install the R package, know the URL of the GRAPLEr Web service, and decide how the simulations should be setup.

5.4 Integration with catchment and climate models

GLM simulations may be coupled with catchment models, such as the Soil Water Assessment Tool (SWAT) or similar catchment models, simply by converting the catchment model output into the inflow file format via conversion scripts. Similarly, scripts exist for coupling GLM with the Weather Research Forecasting (WRF) model, or similar climate models, for specification of the meteorological input file from weather prediction simulations.

The above coupling approaches require the models to be run in sequence, however, for the simulation of lake-wetland-groundwater systems, two-way coupling is required to account for the flow of water into and out of the lake throughout the simulation. For these applications, the interaction can be simulated using GLM coupled with the 3D groundwater flow model, FEFLOW (https://www.mikepoweredbydhi.com/products/feflow). For this case the GLM code is compiled as a Dynamic Link Library (DLL) and loaded into FEFLOW as a plug-in module. The coupling between GLM and FEFLOW is implemented using a one-step lag between the respective solutions of the groundwater and lake models. This approach, in most simulations, does not introduce a significant error, however, error can be assessed and reduced using smaller time step lengths. FEFLOW models can be simulated for flow-only, or including heat and/or solute transport. Depending on the

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

simulation mode, GLM accounts for the different process variables, assigning boundaries for lake level, salinity and temperature accordingly.

The GLM module was designed to accommodate situations of variable lake geometry, by using a dry-lake/wet-lake approach. In this approach, dry-lake areas are defined as those above the current lake level and wet-lake areas as below the current lake level. Different boundary types in FEFLOW are assigned to dry-lake and wet-lake areas (Figure 19). The calibration of such coupled models is often complex, given the large number of parameters and sensitivities when different sources of information are utilised (for example flow and water level measurements). The FEFLOW-GLM coupling structure allows for a relatively straightforward integration with PEST (Doherty, 2015), based on existing FEFLOW workflows.

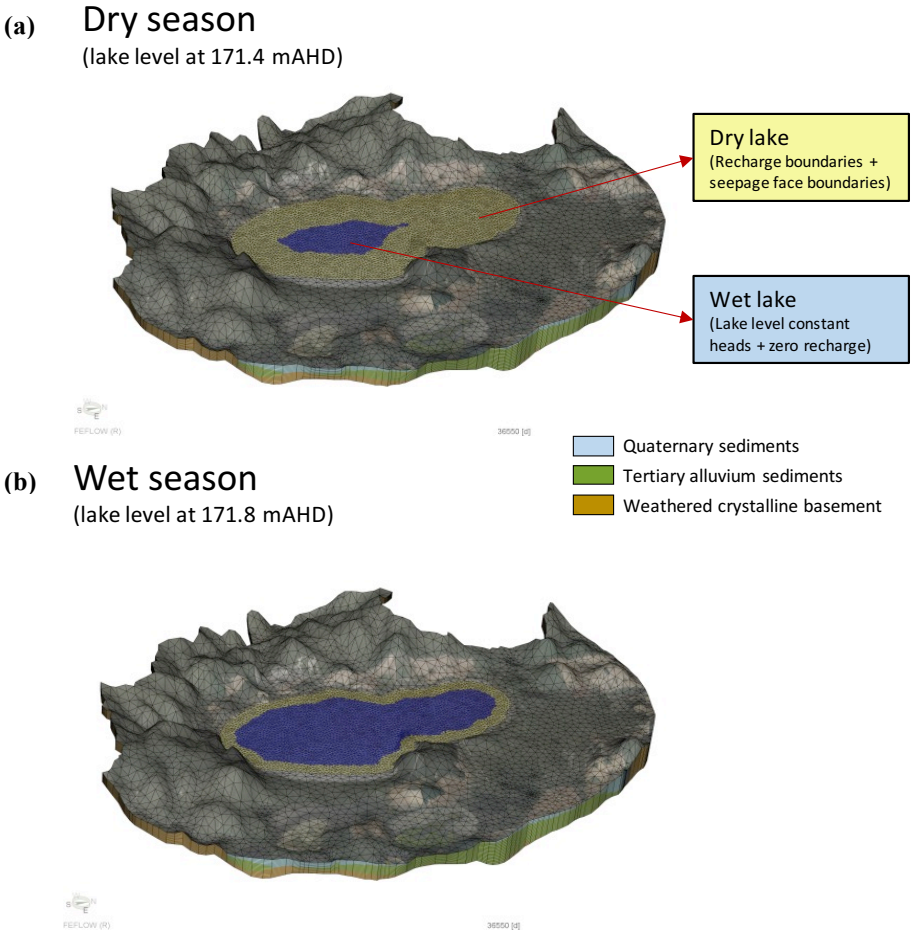

Figure 19: Example of lake boundary changes during wet and dry cycles from Lake Muir, Australia. GLM water level is communicated to FEFLOW to each time-step and used as a constant head boundary condition for all wet cells.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

6 GLM as a tool for teaching environmental science and ecology

Environmental modelling is integral for understanding complex ecosystem responses to anthropogenic and natural drivers, and also provides a valuable tool for engaging students learning environmental science (Carey and Gougis, 2017). Previous pedagogical studies have demonstrated that engaging students in modelling provides cognitive benefits, enabling them to build new scientific knowledge and conceptual development (Stewart et al., 2005; Zohar and Dori 2011). For example, modelling forces students to analyze patterns in data, create evidence-based hypotheses for those patterns and make their hypotheses explicit, and develop predictions of future conditions (Stewart et al., 2005). As a result, the U.S. National Research Council has recently integrated modelling into the *Next Generation Science Standards*, which provide recommendations for primary and secondary school science pedagogy in the United States (NRC, 2013). However, it remains rare for undergraduate and graduate science courses to include the computer-based modelling that environmental scientists need to manage natural ecosystems.

A teaching module for the use of GLM within undergraduate and graduate classrooms has been developed to explore lake responses to climate change (Carey and Gougis, 2017). The GLM module, called the "Climate Change Effects on Lake Temperatures", teaches students how to set up a simulation for a model lake within R. After they are able to successfully run their lake simulations, they force the simulation with climate scenarios of their own design to examine how lakes may change in the future. To improve computational efficiency, students also learn how to submit, retrieve, and analyze hundreds of model simulations through distributed computing overlay networks embedded via the GRAPLEr interface, described above. Hence, students participating in the module learn computing and quantitative skills in addition to improving their understanding of how climate change is affecting lake ecosystems.

Initial experiences teaching GLM as well as pre- and post-assessments indicate that participation in the module improves students' understanding of lake responses to climate change (Carey and Gougis, 2017). By modifying GLM boundary condition data and exploring model output, students are able to better understand the processes that control lake responses to altered climate, and improve their predictions of future lake change. Moreover, the module exposes students to computing and modelling tools not commonly experienced in most university classrooms, building competence with manipulating data files, scripting, creating figures and other visualizations, and statistical and time series analysis; all skills that are transferrable for many other applications. Thus, following previous studies (Schwarz and White 2005, Schwarz et al. 2009), we predict that increased experience with GLM modelling will not only build students' understanding of lake ecosystem concepts but also their ability to interpret natural phenomena and generate new scientific knowledge.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

7 Conclusions

As part of GLEON activities, the emergence of complex questions about how different lake types across the world are responding to climate change and land-use change has created the need for a robust, accessible community code suitable for a diverse range of lake types and simulation contexts. Here, GLM is presented as a tool that meets many of the needs of network participants for their individual lake simulation requirements, in addition to being suitable for application in a distributed way across tens to thousands of lakes for regional and global scale assessments. Recent examples include an application of the model for assessing how the diversity of >2000 lakes in lake-rich landscape in Wisconsin respond to meteorological conditions and projected warming (Read et al., 2014), and given its computationally efficient nature it is envisioned to be made available as a library for use within in land-surface models (e.g., the Community Land Model, CLM), allowing improved representation of lake dynamics in regional hydrological or climate assessments. With further advances in the degree of resolution and scope of earth system models, we further envisage GLM as an option suitable to be embedded within these models to better allow the simulation of lake stratification, air-water interaction of momentum and heat, and also biogeochemically relevant variables associated with contemporary questions about greenhouse gases emissions such as CO_2, CH_4, and N_2O.

Since the model is one-dimensional, it assumes no horizontal variability in the simulated water layers and users must therefore ensure their application of the model is suited to this assumption. For stratified systems, the parameterization of mixing due to internal wave and inflow intrusion dynamics is relatively simple, making the model ideally suited to longer-term investigations ranging from weeks to decades (depending on the domain size), and for coupling with biogeochemical models to explore the role that stratification and vertical mixing play on lake ecosystem dynamics. However, the model can also be used for shallow lakes, ponds and wetland environments where the water column is relatively well mixed. In order to better define the typical level of model performance across these diverse lake types, a companion paper by Bruce et al. (2017), has undertaken a systematic assessment of the model's error structure against 31 lakes from across GLEON, to which readers can refer to for further guidance. In cases where the assumption of one-dimensionality is not met for a particular lake application, then it is possible for users to apply two or three dimensional models; potentially these can be temporally nested within a longer term GLM simulation.

This paper has focused on description of the hydrodynamic model, but we highlight that the model is a platform for coupling with advanced biogeochemical and ecological simulation libraries for water quality prediction and integrated ecosystem assessments. As with most coupled hydrodynamic-ecological modelling platforms, GLM handles the boundary conditions and transport of variables simulated within these libraries, including the effects of inflows, vertical mixing, and evapo-concentration. Whilst the interface to these libraries is straightforward, the Lagrangian approach adopted within GLM for

simulation of the water column necessitates the adoption of sediment zones on a static grid that is independent from the water column numerical grid.

More advanced workflows for operation of the model within distributed computing environments and with data assimilation algorithms is an important application when used within GLEON capabilities related to high frequency data and its interpretation. The 1D nature of the model makes the run-times modest and therefore the model is suitable for application within more intensive parameter identification and uncertainty assessment procedures. This is particularly relevant as the needs for network participants to expand model configurations to further include biogeochemical and ecological state variables. It is envisioned that continued application of the model to lakes of GLEON will allow us to improve parameter estimates and ranges, and this will ultimately support other users of the model in identifying parameter values, and assigning parameter prior distributions. Since many of the users the model is intended for may not have access to the necessary cyberinfrastructure, the use of GLM with the open-source GRAPLEr software in the R environment provides access to otherwise unavailable distributed computing resources. This has the potential to allow non-expert modellers within the science community to apply good modelling practices by automating boundary condition and parameter sensitivity assessments, with technical aspects of simulation management abstracted from the user.

Finally, the role of models in informing and educating members of the network and the next generation of hydrologic and ecosystem modellers has been identified as a critical element of synthesis activities and supporting cross-disciplinary collaboration (Weathers et al., 2017). Initial use of GLM within the classroom has found that teaching modules integrating GLM into classes improves students' understanding of how climate change is altering lake ecosystems.

Code availability

The GLM code is provided as open-source under the GNU GPL3 license, and version controlled via the GitHub repository: https://github.com/AquaticEcoDynamics/GLM.

Data availability

The five example lakes used to demonstrate the model operation are described along with model input files (and associated hydrologic and meteorological forcing data) within the GitHub repository: https://github.com/AquaticEcoDynamics/GLM/tree/master/Examples/2.4.0

Acknowledgments

The primary code of GLM has been developed by MH, LB, CB, BB and DH at The University of Western Australia in collaboration with researchers participating in GLEON and benefited directly from support provided by the NSF Research Coordination Network Award. Whilst GLM is a new code, it is based on the large body of historical research and publications produced by the Centre for Water Research at the University of Western Australia, which we acknowledge for the inspiration, development and testing of several of the model algorithms that have been adopted. Funding for initial development of the GLM code was from the U.S. NSF Cyber-enabled Discovery and Innovation grant awarded to PH (lead investigator) and colleagues from 2009-2014 (NSF CDI-0941510), and subsequent development was supported by the Australian Research Council projects awarded to MH and colleagues (ARC projects LP0990428, LP130100756 and DP130104087). Funding for development of the GLM teaching module and GRAPLEr was sup ported from NSF ACI-1234983 and NSF EF-1702506 awarded to CC. Provision of the environmental symbols used for the GLM scientific diagrams are courtesy of the Integration and Application Network, University of Maryland Center for Environmental Science. Joanne Moo and Aditya Singh also provided support in model setup and testing.

References

Ashton, G.D. (ed.), 1986. River and lake ice engineering. Water Resources Publications, Littleton, Colorado, U.S.A.

Antenucci, J.P., Brookes, J.D. and Hipsey, M.R., 2005. A simple model for quantifying" Cryptosporidium" transport, dilution, and potential risk in reservoirs. *Journal of the American Water Works Association*, 97(1): 86-93.

Babanin, A.V. and Makin, V.K., 2008. Effects of wind trend and gustiness on the sea drag: Lake George study. *Journal of Geophysical Research: Oceans,* 113(C2).

Bird, R.E., 1984. A simple, solar spectral model for direct-normal and diffuse horizontal irradiance. *Solar energy*, 32: 461-471.

Briegleb, B.P., Minnis, P., Ramanathan, V. and Harrison, E., 1986. Comparison of regional clear-sky albedos inferred from satellite observations and model computations. *Journal of Climate and Applied Meteorology*, 25(2): 214-226.

Bruce, L.C., Frassl, M.A., Arhonditsis, G.B., Gal, G., Hamilton, D.P., Hanson, P.C., et al., 2017. A multi-lake comparative analysis of the General Lake Model (GLM): Stress-testing across a global observatory network. *Environmental Modelling & Software*, accepted pending final revision.

Bruggeman, J. and Bolding, K., 2014. A general framework for aquatic biogeochemical models. *Environmental Modelling & Software*, 61: 249-265.

Brutsaert, W., 1975. On a derivable formula for long-wave radiation from clear skies, *Water Resources Research*, 11(5): 742–744.

Bueche, T., Hamilton, D.P. and Vetter, M., 2017. Using the General Lake Model (GLM) to simulate water temperatures and ice cover of a medium-sized lake: a case study of Lake Ammersee, Germany. *Environmental Earth Sciences*, *76*(13), p.461. doi.org/10.1007/s12665-017-6790-7

Businger, J.A., Wyngaard, J.C., Izumi, Y. and Bradley, E.F., 1971. Flux profile relationships in the atmospheric surface layer. *Journal of Atmospheric Sciences*, 28: 181-189.

Carey, C.C. and Gougis, R.D., 2017. Simulation modeling of lakes in undergraduate and graduate classrooms increases comprehension of climate change concepts and experience with computational tools. *Journal of Science Education and Technology*, 26: 1-11.

Chung, E. G., Schladow, S. G., Perez-Losada, J., and Robertson, D. M., 2008. A linked hydrodynamic and water quality model for the Salton Sea. Hydrobiologia 604:57–75.

Chung, S.W., Imberger, J., Hipsey, M.R. and Lee, H.S., 2014. The influence of physical and physiological processes on the spatial heterogeneity of a Microcystis bloom in a stratified reservoir. *Ecological Modelling*, *289*, pp.133-149.

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: 172–185.

Doherty, J., 2015. Calibration and Uncertainty Analysis for Complex Environmental Models. Watermark Numerical Computing, Brisbane, Australia. ISBN: 978-0-9943786-0-6

Dyer, A.J., 1974. A review of flux-profile relationships. *Boundary-Layer Meteorology*, *7*: 363-372.

Fischer, H.B., List, E.G., Koh, R.C.Y., Imberger, J. and Brooks, N.H., 1979. Mixing in Inland and Coastal Waters. Academic Press. New York, NY, 483 pp.

Francey, R.J. and Garratt, J.R., 1978. Eddy flux measurements over the ocean and related transfer coefficients. *Boundary-Layer Meteorology*, 14: 153-166.

Gal, G., Imberger, J., Zohary, T., Antenucci, J., Anis, A. and Rosenberg, T., 2003. Simulating the thermal dynamics of Lake Kinneret. *Ecological Modelling*, 162: 69-86.

Gal, G., Hipsey, M.R., Parparov, A., Wagner, U., Makler, V. and Zohary, T., 2009. Implementation of ecological modeling as an effective management and investigation tool: Lake Kinneret as a case study. Ecological Modelling, 220(13), pp.1697-1718.

Ganguly, A., Agrawal, A., Boykin, P. O. and Figueiredo, R., 2006. IP over P2P: Enabling self-configuring virtual IP networks for grid computing. In International Parallel and Distributed Processing Symposium.

Gill, A.E., 1982. Atmosphere-ocean dynamics (Vol. 30). Academic press.

Gu, R. and Stefan, H.G., 1993. Validation of cold climate lake temperature simulation. *Cold regions science and technology*, 22: 99-104.

Haario, H., Laine, M., Mira, A. and Saksman, E., 2006. DRAM: Efficient adaptive MCMC, *Statistics and Computing,* 16: 339-354.

Hamilton, D.P. and Schladow, S.G. 1997. Water quality in lakes and reservoirs. Part I Model description. *Ecological Modelling*, **96**: 91–110.

Hamilton, D.P., Carey, C.C., Arvola, L., Arzberger, P., Brewer, C., Cole, J.J., Gaiser, E., Hanson, P.C., Ibelings, B.W., Jennings, E. and Kratz, T.K., 2015. A Global Lake Ecological Observatory Network (GLEON) for synthesising high-frequency sensor data for validation of deterministic ecological models. *Inland Waters*, 5(1): 49-56.

Hanson, P.C., Weathers, K.C. and Kratz, T.K., 2016. Networked lake science: how the Global Lake Ecological Observatory Network (GLEON) works to understand, predict, and communicate lake ecosystem response to global change. Inland Waters, 6(4): 543-554.

Harvey, L.D.D., 1990. Testing alternative parameterizations of lateral melting and upward basal heat flux in a thermodynamic sea ice model. *Journal of Geophysical Research,* 95: 7359-7365.

Henderson-Sellers, B., 1986. Calculating the surface energy balance for lake and reservoir modeling: A review. *Reviews of Geophysics*, 24(3): 625-649.

Hicks, B.B., 1975. A procedure for the formulation of bulk transfer coefficients over water. *Boundary Layer Meteorology*, 8: 515-524.

Hicks, B.B., 1972. Some evaluations of drag and bulk transfer coefficients over water. *Boundary Layer Meteorology*, 3: 201:213.

Hipsey, M.R. and Sivapalan, M., 2003. Parameterizing the effect of a wind-shelter on evaporation from small waterbodies. *Water Resources Research*, 39(12): 1339.

Hipsey, M.R., Bruce, L.C. and Hamilton, D.P., 2013. Aquatic EcoDynamics (AED) Model Library Science Manual. The University of Western Australia Technical Manual, Perth, Australia, p.34.

Hipsey, M.R., Hamilton, D.P., Hanson, P.C., Carey, C.C., Coletti, J.Z., Read, J.S., Ibelings, B.W., Valesini, F.J. and Brookes, J.D., 2015. Predicting the resilience and recovery of aquatic systems: A framework for model evolution within environmental observatories. *Water Resources Research*, 51(9): 7023-7043.

Hocking, G.C. and Patterson, J.C., 1991. Quasi-two-dimensional reservoir simulation model. *Journal of Environmental Engineering*, 117: 595-613.

Hu, F., Bolding, K., Bruggeman, J., Jeppesen, E., Flindt, M., van Gerven, L.P.A., Janse, J.H., Janssen, A.B.G., Kuiper, J.J., Mooij, W.M. and Trolle, D. 2016. FABM-PCLake - Linking aquatic ecology with hydrodynamics. Geoscientific Model Development 9: 2271-2278

Idso, S.B. and Jackson, R.D., 1969. Thermal radiation from the atmosphere. *Journal of Geophysical Research*, 74: 5397-5403.

Imberger, J., Patterson, J., Hebbert, B. and Loh, I., 1978. Dynamics of reservoir of medium size. *Journal of the Hydraulics Division - ASCE*, 104 No HY5: 725-743.

Imberger, J. and Patterson, J.C., 1981. *A dynamic reservoir simulation model-DYRESM:5*. In: H.B. Fischer (ed.), Transport Models for Inland and Coastal Waters. Academic Press, New York: 310-361.

Imberger, J. and Patterson, J.C., 1990. *Physical Limnology*. In: T. Wu (ed.), Advances in applied mechanics 27. Academic Press. Boston. U.S.A.

Imboden, D.M. and Wüest, A., 1995. *Mixing Mechanisms in Lakes*, p. 83-138. In: A. Lerman, D.M. Imboden and J.R. Gat (eds.), Physics and Chemistry of Lakes. Springer-Verlag.

Janssen, A.B.G., Arhonditsis, G.B., Beusen, A., Bolding, K., Bruce, L., Bruggeman, J., Couture, R.M., Downing, A.S., Elliott, J.A., Frassl, M.A., Gal, G., Gerla, D.J., Hipsey, M.R., Hu, F., Ives, S.C., Janse, J., Jeppesen, E., Jöhnk, K.D., Kneis, D., Kong, X., Kuiper, J.K., Lehmann, M., Lemmen, C., Ozkundakci, D., Petzoldt, T., Rinke, K., Robson, B.J., Sachse, R., Schep, S., Schmid, M., Scholten, H., Teurlincx, S., Trolle, D., Troost, T.A., Van Dam, A., Van Gerven, L.A., Weijerman, M., Wells S.A. and Mooij, W.M., 2015. Exploring, exploiting and evolving diversity of aquatic ecosystem models: a community perspective. *Aquatic ecology*, 49(4): 513-548.

Jellison, R. and Melack, J.M., 1993. Meromixis and vertical diffusivities in hypersaline Mono Lake, California. *Limnology and Oceanography*, **38**: 1008–1019.

Ji, Z.G., 2008. Hydrodynamics and water quality: modeling rivers, lakes, and estuaries. John Wiley & Sons.

Kara, E.L., Hanson, P., Hamilton, D., Hipsey, M.R., McMahon, K.D., Read, J.S., Winslow, L., Dedrick, J., Rose, K., Carey, C.C. and Bertilsson, S., 2012. Time-scale dependence in numerical simulations: assessment of physical, chemical, and

biological predictions in a stratified lake at temporal scales of hours to months. *Environmental Modelling & Software*, *35*, pp.104-121.

Kim, J-W., 1976. A generalized bulk model of the oceanic mixed layer. *Journal of Physical Oceanography*, 6: 686-695.

Kirk, J.T.O., 1994. Light and photosynthesis in aquatic ecosystems. Cambridge University Press.

Kirillin, G., Hochschild, J., Mironov, D., Terzhevik, A., Golosov, S. and Nützmann, G., 2011. FLake-Global: Online lake model with worldwide coverage. Environmental Modelling & Software, 26(5), pp.683-684.

Kleinhans, M.G. and Grasmeijer, B.T., 2006. Bed load transport on the shoreface by currents and waves. *Coastal Engineering*, 53: 983-996.

Klug, J.L., Richardson, D.C., Ewing, H.A., Hargreaves, B.R., Samal, N.R., Vachon, D., Pierson, D.C., Lindsey, A.M., O'Donnell, D.M., Effler, S.W. and Weathers, K.C., 2012. Ecosystem Effects of a Tropical Cyclone on a Network of Lakes in Northeastern North America. *Environmental Science & Technology*, 46 (21): 11693-11701.

Kraus, E.B. and Turner, J.S., 1967. A one-dimensional model of the seasonal thermocline: II The general theory and its consequences. *Tellus*, 19: 98-106.

Laenen A. and LeTourneau A.P., 1996. Upper Klamath Lake nutrient loading study – Estimate of wind-induced resuspension of bed sediment during periods of low lake elevation. *US Geological Survey Open-File Report, 95-414*, 11 pp

Launiainen, J., 1995. Derivation of the relationship between the Obukhov stability parameter and the bulk Richardson number for flux-profile studies. *Boundary Layer Meteorology*, **76**: 165-179.

Launiainen, J. and Cheng, B., 1998. Modelling of ice thermodynamics in natural water bodies. *Cold Region Science and Technology*, **27**: 153-178.

Launiainen, J. and Vihma, T., 1990. Derivation of turbulent surface fluxes—An iterative flux-profile method allowing arbitrary observing heights. *Environmental Software*, 5: 113-124.

Le Roux, J.P., 2007. A simple method to determine breaker height and depth for different deepwater wave height/length ratios and sea floor slopes. *Coastal Engineering*, 54: 271-277.

Luo, L., Hamilton, D.P. and Han, B., 2010. Estimation of total cloud cover from solar radiation observations at Lake Rotorua, New Zealand. *Solar Energy*, 84: 501-506.

Magee, M.R., Wu, C.H., Robertson, D.M., Lathrop, R.C. and Hamilton, D.P., 2016. Trends and abrupt changes in 104 years of ice cover and water temperature in a dimictic lake in response to air temperature, wind speed, and water clarity drivers. Hydrology and Earth System Sciences, 20(5), p.1681.

Makler-Pick, V., Gal, G., Shapiro, J. and Hipsey, M.R., 2011. Exploring the role of fish in a lake ecosystem (Lake Kinneret, Israel) by coupling an individual-based fish population model to a dynamic ecosystem model. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(7), pp.1265-1284.

Markfort, C.D., Perez, A.L.S., Thill, J. W., Jaster, D.A., Porté-Agel, F. and Stefan, H.G. 2010. Wind sheltering of a lake by a tree canopy or bluff topography. *Water Resources Research*, 46: 1–13.

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, Series A Dynamic Meteorology And Oceanography*, 64: 1-22.

Matzinger, A., Schmid, M., Veljanoska-Sarafiloska, E., Patceva, S., Guseska, D., Wagner, B., Müller, B., Sturm, M. and Wüest, A., 2007. Eutrophication of ancient Lake Ohrid: global warming amplifies detrimental effects of increased nutrient inputs, *Limnology and Oceanography*, 52(1), 338-353, doi:10.4319/lo.2007.52.1.0338.

McKay, G.A., 1968. Problems of measuring and evaluating snow cover. In: *Proceedings Workshop Seminar of Snow Hydrology.(Secretariat Canadian National Committee for the IHD*, Ottawa: 49-62.

McCord, S.A. and Schladow, S.G., 1998. Numerical simulations of degassing scenarios for CO2-rich Lake Nyos, Cameroon. *Journal of Geophysical Research: Solid Earth*, 103(B6): 12355-12364.

Menció, A., Casamitjana, X., Mas-Pla, J., Coll, N., Compte, J., Martinoy, M., Pascual, J. and Quintana, X.D., 2017. Groundwater dependence of coastal lagoons: The case of La Pletera salt marshes (NE Catalonia). Journal of Hydrology, 552, pp.793-806.

Mueller H, Hamilton DP and Doole GJ. 2016. Evaluating services and damage costs of degradation of a major lake ecosystem. Ecosystem Services 22: 370-380

Mooij, W.M., Trolle, D., Jeppesen, E., Arhonditsis, G., Belolipetsky, P.V., Chitamwebwa, D.B.R., Degermendzhy, A.G., DeAngelis, D.L., De Senerpont Domis, L.N., Downing, A.S., Elliott, A.E., Fragoso Jr, C.R., Gaedke, U., Genova, S.N., Gulati, R.D., Håkanson, L., Hamilton, D.P., Hipsey, M.R., Hoen, J., Hülsmann, S., Los, F.J., Makler-Pick, V., Petzoldt, T., Prokopkin, I.G., Rinke, K., Schep, S.A., Tominaga, K., Van Dam, A.A., Van Nes, E.H., Wells, S.A. and Janse, J.H., 2010. Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquatic Ecology*, **44**: 633–667.

Monin, A.S. and Obukhov, A.M., 1954. Basic laws of turbulent mixing in the atmosphere near the ground. *Jr. Akad. Nauk SSSR Geofiz. Inst.,* **24**: 163-187.

National Research Council (NRC), 2013. Next generation science standards: For states, by states. Washington, DC: The National Academies Press.

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., et al., 2015. Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, 42(24): 10,773–10,781.

Patterson, J.C., Hamblin, P.F. and Imberger, J., 1984. Classification and dynamics simulation of the vertical density structure of lakes. *Limnology and Oceanography*, 29: 845-861.

Patterson, J.C. and Hamblin, P.F., 1988. Thermal simulation of a lake with winter ice cover. *Limnology and Oceanography*, 33: 323-338.

Paulson, C. A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of Applied Meteorology*, 9: 857-861.

Peeters, F., Straile, D. Loke, A. and Livingstone, D. M., 2007. Earlier onset of the spring phytoplankton bloom in lakes of the temperate zone in a warmer climate. *Global Change Biology*. 13: 1898–1909, doi:10.1111/j.1365-2486.2007.01412.x

Perroud, M., Goyette, S., Martynov, A., Beniston, M. and Annevillec, O., 2009. Simulation of multiannual thermal profiles in deep Lake Geneva: A comparison of one-dimensional lake models. *Limnology and Oceanography*, *54*(5), pp.1574-1594.

Porter, J.H., Hanson, P.C. and Lin, C.C., 2012. Staying afloat in the sensor data deluge. *Trends In Ecology & Evolution*, 27(2): 121-129.

Read, J.S., Hamilton, D.P., Jones, I.D., Muraoka, K., Winslow, L.A., Kroiss, R., Wu, C.H. and Gaiser, E., 2011. Derivation of lake mixing and stratification indices from high-resolution lake buoy data. *Environmental Modelling & Software,* **26**: 1325–1336.

Read, J.S., Hansen, G., Van Den Hoek, J., Hanson, P.C., Bruce, L.C. and Markfort, C.D., 2014. Simulating 2368 temperate lakes reveals weak coherence in stratification phenology. *Ecological Modelling*, 291: 142-150.

Read, J.S., Gries, C., Read, E.K., Klug, J., Hanson, P.C., Hipsey, M.R., Jennings, E., O'Reilly, C., Winslow, L., Pierson, D., McBride, C. and Hamilton, D.P., 2016. Generating community-built tools for data sharing and analysis in environmental networks. *Inland Waters*, 6(4): 637-644.

Rigosi, A., Hanson, P.C., Hamilton, D.P., Hipsey, M., Rusak, J. A., Bois, J., Sparber, K., Chorus, I., Watkinson, A.J., Qin, B., Kim, B. and Brookes, J.D., 2015. Determining the probability of cyanobacterial blooms: the application of Bayesian networks in multiple lake systems. *Ecological Applications*, 25: 186–199.

Rogers, C.K., Lawrence, G.A. and Hamblin, P.F., 1995. Observations and numerical simulation of a shallow ice-covered mid-latitude lake. *Limnology and Oceanography,* **40**: 374-385.

Romarheim, A.T., Tominaga, K., Riise, G. and Andersen, T., 2015. The importance of year-to-year variation in meteorological and runoff forcing for water quality of a temperate, dimictic lake. *Hydrology and Earth System Sciences*, 19(6): 2649-2662.

Salmon, S.U., Hipsey, M.R., Wake, G.W., Ivey, G.N. and Oldham, C.E., 2017. Quantifying Lake Water Quality Evolution: Coupled Geochemistry, Hydrodynamics, and Aquatic Ecology in an Acidic Pit Lake. Environmental Science & Technology, 51(17): 9864-9875.

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), pp.45-60.

Schwarz, C.V. and White, B.Y., 2005. Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23(2): 165-205.

Schwarz, C.V., Reiser, B.J., Davis, E.A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B. and Krajcik, J., 2009. Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6): 632-654.

Sherman, F.S., Imberger, J. and Corcos, G.M., 1978. Turbulence and mixing in stably stratified waters. *Annual Review of Fluid Mechanics,* 10: 267-288.

Snortheim, C.A., P.C. Hanson, K.D. McMahon, J.S. Read, C.C. Carey, and H.A. Dugan. Meteorological drivers of hypolimnetic anoxia in a eutrophic, north temperate lake. Ecological Modelling. 343: 39-53. DOI: 10.1016/j.ecolmodel.2016.10.014

Spigel, B., 1978. Wind mixing in Lakes. PhD thesis, University of California, Berkeley.

Spigel, R.H., Imberger, J. and Rayner, K.N., 1986. Modeling the diurnal mixed layer. *Limnology and Oceanography*, 31: 533-556.

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), pp.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), pp.1337-1352.

Stewart, J., Cartier, J.L. and Passmore, C.M., 2005. Developing understanding through model-based inquiry. In M.S. Donovan & J.D. Bransford (Eds.), How students learn (pp. 515–565). Washington, DC: National Research Council.

Strub, P.T. and Powell, T.M., 1987. Surface temperature and transport in Lake Tahoe: inferences from satellite (AVHRR) imagery. *Continental Shelf Research,* 7: 1001-1013.

Subratie, K., Aditya, S., Figueiredo, R., Carey, C.C. and Hanson, P.C., 2017. GRAPLEr: A distributed collaborative environment for lake ecosystem modeling that integrates overlay networks, high-throughput computing, and web services. *Concurrency and Computation: Practice and Experience*, 29(13): e4139.

Swinbank, W.C., 1963. Longwave radiation from clear skies. *Quarterly Journal of the Royal Meteorological Society*, 89: 339-348.

Tennessee Valley Authority (TVA), 1972. Heat and mass transfer between a water surface and the atmosphere. Water Resources Research Laboratory Report 14, Report No. 0-6803.

Tabata, S., 1973. A simple but accurate formula for the saturation vapour pressure over liquid water. *Journal of Applied Meteorology*, 12: 1410-1411.

Thain, D., Tannenbaum, T. and Livny, M. (2005). "Distributed Computing in Practice: the Condor Experience" (PDF). *Concurrency and Computation: Practice and Experience*, 17 (2–4): 323–356.

Ticehurst J.L., Newham LTH, Rissik D, Letcher RA and Jakeman AJ. 2007. A Bayesian network approach for assessing the sustainability of coastal lakes in New South Wales, Australia. Environmental Modelling & Software 22(8):1129-1139

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. and Kortelainen, P.L., 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, 54(6part2): 2298-2314.

Trolle, D., Hamilton, D.P., Hipsey, M.R., Bolding, K., Bruggeman, J., Mooij, W. M., Janse, J. H., Nielsen, A., Jeppesen, E., Elliott, J.E., Makler-Pick, V., Petzoldt, T., Rinke, K., Flindt, M. R., Arhonditsis, G.B., Gal, G., Bjerring, R., Tominaga, K., Hoen, J., Downing, A.S., Marques, D. M., Fragoso Jr, C.R., Søndergaard, M. and Hanson, P.C., 2012. A community-based framework for aquatic ecosystem models. *Hydrobiologia*, 683: 25-34.

UNESCO, 1981. Technical papers in Marine Science. No. 36.

Vavrus, S.J., Wynne, R.H. and Foley, J.A., 1996. Measuring the sensitivity of southern Wisconsin lake ice to climate variations and lake depth using a numerical model. *Limnology and Oceanography*, 41: 822-831.

Vickers, D., Mahrt, L. and Andreas, E.L., 2013. Estimates of the 10-m neutral sea surface drag coefficient from aircraft eddy-covariance measurements. *Journal of Physical Oceanography*, 43: 301-310.

Weathers, K.C., Groffman, P.M., Van Dolah, E., Bernhardt, E.S., Grimm, N.B., McMahon, K.D., Schimel, J., Paolisso, M., Maranger, R.J., Baer, S., Brauman, K.A. and Hinckley, E., 2016. Frontiers in Ecosystem Ecology from a Community Perspective: The Future is Boundless and Bright. *Ecosystems*, 19(5): 753-770.

Weber, M., Rinke, K., Hipsey, M.R. and Boehrer, B., 2017. Optimizing withdrawal from drinking water reservoirs to reduce downstream temperature pollution and reservoir hypoxia. *Journal of Environmental Management*, 197: 96-105.

Weinstock, J., 1981. Vertical turbulence diffusivity for weak or strong stable stratification. *Journal of Geophysical Research*, 86(C10): 9925-9928.

Williamson, C.E., Saros, J.E., Vincent, W.F. and Smol, J.P., 2009. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography*, 54(6part2), 2273-2282.

Woolway, R.I., Verburg, P., Merchant, C.J., Lenters, J.D., Hamilton, D.P., Brookes, J., Kelly, S., Hook, S., Laas, A., Pierson, D. and Rimmer, A., 2017. Latitude and lake size are important predictors of over-lake atmospheric stability. *Geophysical Research Letters*, 44 (17), 8875–8883. DOI:10.1002/2017GL073941

Wu J., 1973. Wind induced entrainment across a stable density interface. *Journal of Fluid Mechanics*, 61: 275-278.

Xenopoulos, M.A. and Schindler, D.W., 2001. The environmental control of near-surface thermoclines in boreal lakes. *Ecosystems*, 4: 699-707.

Yajima, H. and Yamamoto, S., 2015. Improvements of radiation estimations for a simulation of water temperature in a reservoir. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering*, 71(4): 775-780.

Yao, H., Samal, N.R., Joehnk, K.D., Fang, X., Bruce, L.C., Pierson, D.C., Rusak, J.A. and James, A., 2014. Comparing ice and temperature simulations by four dynamic lake models in Harp Lake: past performance and future predictions. *Hydrological Processes*, 28: 4587-4601.

Yeates, P.S. and Imberger, J., 2003. Pseudo two-dimensional simulations of internal and boundary fluxes In stratified lakes and reservoirs. *International Journal of River Basin Research,* **1**: 1-23.

Zhang, W. and Arhonditsis G.B., 2009. A Bayesian hierarchical framework for calibrating aquatic biogeochemical models, *Ecological Modelling*, 220(18): 2142-2161.

Zohar, A. and Dori, Y.J., 2012. *Introduction*. In: Metacognition in science education (pp. 1-19). Springer Netherlands.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

Table 1. Summary of GLM physical parameters with recommended values and references.

Symbol	glm.nml ID	Description	Units	Default	Reference	Comments
Model Structure						
h_{min}	min_layer_thick	Minimum layer thickness	m	0.5	Bruce et al. (2017) Bueche et al. (2017)	Standardised for multi-lake comparison Should be estimated relative to lake depth.
h_{max}	max_layer_thick	Maximum layer thickness	m	1.5		
Lake Properties						
K_w	Kw	Extinction coefficient for PAR radiation	m^{-1}	0.2	Lake specific	Should be measured, e.g. mean of simulation period. Can be estimated from Secchi depth.
A_C	critical_area	Critical area below which wind sheltering may occur	m^2	10^7	Xenopoulos and Schindler (2001)	
Surface Thermodynamics						
C_H	ch	Bulk aerodynamic coefficient for sensible heat transfer	-	0.0013	Fischer et al. (1979) Bruce et al. (2017) Bueche et al. (2017)	From Hicks' (1972) collation of ocean and lake data; many studies since use similar values. Internally calculated if atmos stability correction is on.
C_E	ce	Bulk aerodynamic coefficient for latent heat transfer	-	0.0013		
C_M	cd	Bulk aerodynamic coefficient for transfer of momentum	-	0.0013		
λ	-	Latent heat of evaporation	$J\ kg^{-1}$	$2.453x10^6$	Standard	Not adjustable in glm.nml
ε_a	-	Emissivity of the water surface	-	0.985	Standard	Water only, no ice Ice or snow
σ	-	Stefan-Boltzmann constant	$W\ m^{-2}\ K^{-4}$	$5.67x10^{-8}$	Constant	Not adjustable in glm.nml
Mixing Parameters						
C_K	coef_mix_conv	Mixing efficiency - convective overturn	-	0.2	Yeates & Imberger (2003)	Selected from a range given in Spigel et al. (1986)
C_W	coef_wind_stir	Mixing efficiency - wind stirring	-	0.23	Spigel et al. (1986)	From Wu (1973)
C_S	coef_mix_shear	Mixing efficiency - shear production	-	0.3	Sherman et al. (1978)	Best fit of experiments reviewed
C_T	coef_mix_turb	Mixing efficiency - unsteady turbulence (acceleration)	-	0.51	Bruce et al. (2017) Bueche et al. (2017)	
C_{KH}	coef_mix_KH	Mixing efficiency - Kelvin-Helmholtz turbulent billows	-	0.3	Sherman et al. (1978)	"a good rule of thumb..."
C_{HYP}	coef_mix_hyp	Mixing efficiency of hypolimnetic turbulence	-	0.5	Weinstock (1981)	General diffusivities in Jellison and Melack (1993)
Inflows & Outflows						
C_D	strmbd_drag	streambed_drag	-	0.016	Site specific	Set based on inflow stream type
G	seepage_rate	Seepage rate	$m\ day^{-1}$	0		

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

Symbol	glm.nml ID	Description	Units	Default	Reference	Comments
Snow & Ice						
K_{w1}	-	Waveband 1, snow ice light extinction	m^{-1}	48.0	Rogers et al., (1995), Patterson and Hamblin (1988) Ashton (1986) Yao et al., (2014)	
K_{w2}	-	Waveband 2, snow ice light extinction	m^{-1}	20.0		
K_{b1}	-	Waveband 1, blue ice light extinction	m^{-1}	1.5		
K_{b2}	-	Waveband 2, blue ice light extinction	m^{-1}	20.0		
K_{s1}	-	Waveband 1, snow light extinction	m^{-1}	6		
K_{s2}	-	Waveband 2, snow light extinction	m^{-1}	20		
D_z	-	Distance of hear transfer, ice water	m	0.039		
ρ_{white}	-	Density, snow ice	kg m^{-3}	890		
ρ_{blue}	-	Density, blue ice	kg m^{-3}	917		
ρ_{snow}	-	Density, snow	kg m^{-3}	Variable		
c_{pi}	-	Heat capacity, ice	kJ kg^{-1} $^oC^{-1}$	2.1		
c_{pw}	-	Heat capacity, ice	kJ kg^{-1} $^oC^{-1}$	4.2		
K_c	-	Compaction coefficient	-	Variable		
K_m	-	Thermal conductivity, snow ice	W m^{-1} $^oC^{-1}$	2.0		
K_m	-	Thermal conductivity, blue ice	W m^{-1} $^oC^{-1}$	2.3		
K_m	-	Thermal conductivity, snow	W m^{-1} $^oC^{-1}$	Variable		
K_m	-	Thermal conductivity, sediment	W m^{-1} $^oC^{-1}$	1.2		
K_m	-	Thermal conductivity, water	W m^{-1} $^oC^{-1}$	0.57		
L	-	Latent heat of fusion	kJ kg^{-1}	0334		
Bottom Stress						
D	-	Typical particle diameter	m			

5

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Appendix A: Bird solar radiation model

The Bird Clear Sky Model (BCSM) was developed by (Bird, 1984) to predict clear-sky direct beam, hemispherical diffuse, and total hemispherical broadband solar radiation on a horizontal surface. Average solar radiation is computed at the model time-step (e.g., hourly) based on ten user-specified input parameters (Table A1).

Table A1: Parameters required for the BCSM model.

Variable	Description	Example values (e.g., Luo et al., 2010)
Lat	Latitude (degrees, + for N)	-31.77
Long	Longitude (degrees + for E)	116.03
TZ	Time Zone indicated by number of hours from GMT	+7.5
AP	Atmospheric Pressure (millibars)	1013
Oz	Ozone Conc. (atm-cm)	0.279 - 0.324
W	Total Precipitable Water Vapour (atm-cm)	1.1 - 2.2
AOD_{500}	Aerosol Optical Depth at 500 nm	0.033 – 0.1
AOD_{380}	Aerosol Optical Depth at 380 nm	0.038 – 0.15
α_{SW}	Surface albedo	0.2

The solar constant in the model is taken as 1367 W/m^2. This is corrected due to the elliptical nature of the earth's orbit and consequent change in distance to the sun. This calculation gives us the Extra-Terrestrial Radiation ($\hat{\phi}_{ETR}$), at the top of the atmosphere:

$$\hat{\phi}_{ETR} = 1367\left(1.00011 + 0.034221\cos(\Phi_{day}) + 0.00128\sin(\Phi_{day}) + 0.000719\cos(\Phi_{day})\right) \quad \text{(A1)}$$

where the day angle, Φ_{day}, is computed using, *d,* the day number:

$$\Phi_{day} = 2\pi\left(\frac{d-1}{365}\right) \quad \text{(A2)}$$

The solar declination, Φ_{dec} (radians), is computed from:

$$\Phi_{dec} = \left[\begin{array}{c} 0.006918 - 0.399912\,cos(\Phi_{day}) + 0.070257\,sin(\Phi_{day}) - 0.006758\,cos\left(2(\Phi_{day})\right) + \\ 0.000907\,sin(2\Phi_{day}) - 0.002697\,cos\left(3(\Phi_{day})\right) + 0.00148\,sin\left(3(\Phi_{day})\right) \end{array}\right] \quad \text{(A3)}$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

We then solve the equation of time:

$$EQT = \begin{bmatrix} 0.0000075 + 0.001868\ cos(\Phi_{day}) - 0.032077\ sin(\Phi_{day}) \\ -0.014615\ cos\left(2(\Phi_{day})\right) - 0.040849\ sin\left(2(\Phi_{day})\right) \end{bmatrix} \times 229.18 \quad \text{(A4)}$$

in order to compute the hour angle, Φ_{hr}, calculated with noon zero and morning positive as:

$$\Phi_{hr} = 15(hr - 12.5) + Long - 15\,TZ + \left(\frac{EQT}{4}\right) \quad \text{(A5)}$$

where TZ is the time-zone shift from GMT. The zenith angle, Φ_{zen} (radians), is calculated from:

$$\cos(\Phi_{zen}) = \cos(\Phi_{dec})\cos(\Phi_{hr})\cos(Lat) + \sin(\Phi_{dec})\sin(Lat) \quad \text{(A6)}$$

When Φ_{zen} is less than 90°, the air mass factor is calculated as:

$$\mathrm{AM} = \left[cos(\Phi_{zen}) + \frac{0.15}{(93.885 - \Phi_{zen})^{1.25}}\right]^{-1} \quad \text{(A7)}$$

which is corrected for atmospheric pressure, p (hPa),

$$AM_p = \frac{AM\ p}{1013} \quad \text{(A8)}$$

AM_P is then used to calculate the Rayleigh Scattering as:

$$T_{rayleigh} = e^{[(-0.0903\ AM_p{}^{0.84}) + (1 + AM_p - AM_p{}^{1.01})]} \quad \text{(A9)}$$

The effect of ozone scattering is calculated by computing ozone mass, which for positive air mass is:

$$T_{ozone} = \left[1 - \left(0.1611\ (Oz\ AM)\ \left(1 + 139.48\ (Oz\ AM)\right)^{-0.3035}\right) - \frac{0.002715\ (Oz\ AM)}{1 + 0.044\ (Oz\ AM) + 0.0003\ (Oz\ AM)^2}\right] \quad \text{(A10)}$$

The scattering due to mixed gases for positive air mass is calculated as:

$$T_{mix} = e^{[-0.0127\ AMp^{0.26}]} \quad \text{(A11)}$$

Then the water scattering is calculated by getting the water mass:

$$Wm = WAM_p \quad \text{(A12)}$$

where W is the precipitable water vapour. This can be approximated from dew point temperature, eg.:

$$\ln W = a\,T_d + b \quad \text{(A13)}$$

where a and b are regression coefficients which have been taken as 0.09, 0.07, 0.07 and 0.08 for values of a while b is 1.88, 2.11, 2.12 and 2.01 in spring, summer, autumn and winter (Luo et al., 2010).

Then the water scattering effect is calculated as:

$$T_{water} = \left[1 - \frac{(2.4959\ Wm)}{1 + (79.034\ Wm)^{0.6828} + 6.385\ Wm}\right] \quad \text{(A14)}$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

The scattering due to aerosols requires the Aerosol Optical Depth at 380 nm and 500 nm:

$$TauA = 0.2758\, AOD_{380} + 0.35\, AOD_{500} \quad \text{(A15)}$$

and the scattering due to aerosols is then calculated as:

$$T_{aerosol} = e^{(-TauA)^{0.873}\,(1+TauA-TauA^{0.7088})\,AM^{0.9108}} \quad \text{(A16)}$$

We also define:

$$T_{aa} = 1 - [0.1\,(1 - AM + AM^{1.06})\,(1 - T_{aerosol})] \quad \text{(A17)}$$

and:

$$\frac{0.5(1 - T_{rayleigh}) + 0.84(1 - T_{as})}{1 - AM + AM^{1.02}} \quad \text{(A18)}$$

where the 0.84 value used is actually the proportion of scattered radiation reflected in the same direction as incoming radiation.

The direct beam radiation on a horizontal surface at ground level on a clear day is given by,

$$\hat{\phi}_{DB} = 0.9662\, \hat{\phi}_{ETR}\, T_{rayleigh}\, T_{ozone}\, T_{mix}\, T_{watvap}\, T_{aerosol}\, cos(\Phi_{zen}) \quad \text{(A19)}$$

$$\hat{\phi}_{AS} = 0.79\, \hat{\phi}_{ETR}\, T_{ozone} T_{mix}\, T_{watvap}\, T_{aa}\, cos(\Phi_{zen}) \quad \text{(A20)}$$

The total irradiance hitting the surface is therefore (W m^{-2}):

$$\hat{\phi}_{SW} = \frac{\hat{\phi}_{DB} + \hat{\phi}_{AS}}{1 - (\alpha_{SW}\, \alpha_{SKY})} \quad \text{(A21)}$$

The albedo is computed for the sky as:

$$\alpha_{SKY} = 0.068 + (1 - 0.84)\left(1 - \frac{T_{aerosol}}{T_{aa}}\right) \quad \text{(A22)}$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

Appendix B: Non-neutral bulk transfer coefficients

The iterative procedure used in this analysis is conceptually similar to the methodology discussed in detail in Launiainen and Vihma (1990). The first estimate for the neutral drag coefficient is specified as a function of windspeed as it is has been commonly observed that C_{DN} increases with U_{10} (Figure A1). This is modelled by first by estimating:

$$C_{DN-10} = \begin{cases} 0.001 & U_{10} \le 5 \\ 0.001\,(1 + 0.07[U_{10} - 5]) & U_{10} > 5 \end{cases} \quad \text{Option 1 : Francey and Garratt (1978), Hicks (1972)} \tag{A23}$$

$$C_{DN-10} = 1.92\times10^{-7} U_{10}^{3} + 0.00096 \quad \text{Option 2 : Babanin and Makin (2008)}$$

and then computing the Charnock formula with the smooth flow transition (e.g., Vickers et al., 2013):

$$z_o = \frac{\alpha u_*^2}{g} + 0.11\frac{\nu}{u_*} \tag{A24}$$

where α is the Charnock constant (0.012), u_* is the friction velocity ($\sqrt{C_{DN-10}\,U_{10}^2}$) using Eq A23, and the final drag is re-computed using:

$$C_{DN-10} = \left[\frac{k}{\ln\left(\frac{10}{z_o}\right)}\right]^2 \tag{A25}$$

where k is the von Karman constant. Note the neutral humidity/temperature coefficient, C_{HWN-10}, is held constant at the user defined C_H value and doesn't scale with wind speed.

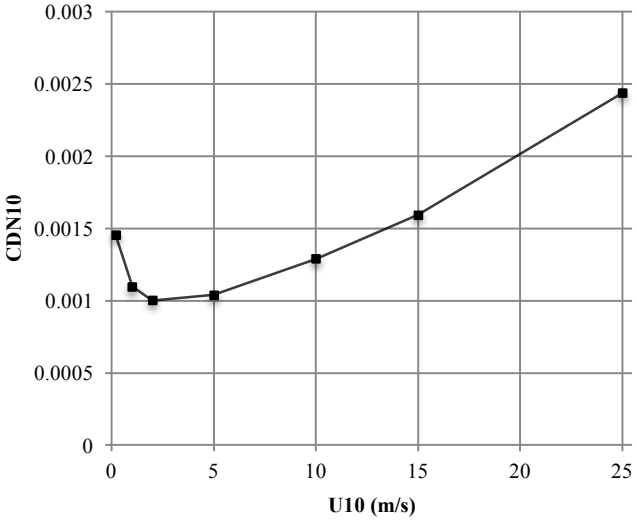

Figure A1: Scaling of the 10m neutral drag coefficient with wind speed (Eqns A23-25)

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Under non-neutral conditions in the atmospheric boundary layer, the transfer coefficients vary due to stratification seen in the air column, as was parameterised by Monin and Obukhov (1954) using the now well-known stability parameter, *z/L*, where *L* is the Obukhov length defined as:

$$L = \frac{-\rho_a u_*^3 \theta_V}{kg\left(\frac{H}{c_p} + 0.61\frac{\theta E}{\lambda}\right)} \tag{A26}$$

where $\theta_V = \theta(1 + 0.61q)$ is the virtual temperature and H and E are the bulk fluxes. Paulson (1970) presented a solution for the vertical profiles of wind speed, temperature and moisture in the developing boundary layer as a function of the Monin-Obukhov stability parameter; the so-called flux-profile relationships:

$$U_z = \frac{u_*}{k}\left[\ln\left(\frac{z}{z_o}\right) - \psi_M\left(\frac{z}{L}\right)\right] \tag{A27a}$$

$$\theta_z - \theta_s = \frac{\theta_*}{k}\left[\ln\left(\frac{z}{z_\theta}\right) - \psi_H\left(\frac{z}{L}\right)\right] \tag{A27b}$$

$$q_z - q_s = \frac{q_*}{k}\left[\ln\left(\frac{z}{z_q}\right) - \psi_E\left(\frac{z}{L}\right)\right] \tag{A27c}$$

where ψ_M, ψ_H and ψ_E are the similarity functions for momentum, heat and moisture respectively, and z_o, z_θ and z_q are their respective roughness lengths. For unstable conditions ($L<0$), the stability functions are defined as (Paulson 1970; Businger *et al.*, 1971; Dyer, 1974):

$$\psi_M = 2\ln\left(\frac{1+x}{2}\right) + \ln\left(\frac{1+x^2}{2}\right) - 2\tan^{-1}x + \frac{\pi}{2} \tag{A28a}$$

$$\psi_E = \psi_H = 2\ln\left(\frac{1+x^2}{2}\right) \tag{A28b}$$

where

$$x = \left[1 - 16\left(\frac{z}{L}\right)^{1/4}\right] \tag{A29}$$

During stable stratification ($L>0$) they take the form:

$$\psi_M = \psi_E = \psi_H = \begin{cases} -5\left(\frac{z}{L}\right) & 0 < \frac{z}{L} < 0.5 \\ 0.5\left(\frac{z}{L}\right)^{-2} - 4.25\left(\frac{z}{L}\right)^{-1} - 7\left(\frac{z}{L}\right) - 0.852 & 0.5 < \frac{z}{L} < 10 \\ \ln\left(\frac{z}{L}\right) - 0.76\left(\frac{z}{L}\right) - 12.093 & \frac{z}{L} > 10 \end{cases} \tag{A30}$$

Substituting equations (17)-(18) into (A27) and ignoring the similarity functions leaves us with neutral transfer coefficients as a function of the roughness lengths:

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

Geoscientific Model Development Discussions

$$C_{XN} = k^2 \left[\ln\left(\frac{z}{z_o}\right)\right]^{-1} \left[\ln\left(\frac{z}{z_X}\right)\right]^{-1} \quad \text{(A31)}$$

where N denotes the neutral value and X signifies either D, H or E for the transfer coefficient and o, θ or q for the roughness length scale. Inclusion of the stability functions into the substitution and some manipulation (Imberger and Patterson, 1990; Launianen and Vihma, 1990) yields the transfer coefficients relative to these neutral values:

$$\frac{C_X}{C_{XN}} = \left[1 + \frac{C_{XN}}{k^2}\left(\psi_M \psi_X - \frac{k\psi_X}{\sqrt{C_{DN}}} - \frac{k\psi_M\sqrt{C_{DN}}}{C_{XN}}\right)\right] \quad \text{(A32)}$$

Hicks (1975) and Launianen and Vihma (1990) suggested an iterative procedure to solve for the stability corrected transfer coefficient using (A32) based on some initial estimate of the neutral value. The surface flux is subsequently estimated according to (17-18) and used to provide an initial estimate for L (equation A26). The partially corrected transfer coefficient is then recalculated and so the cycle goes. Strub and Powell (1987) and Launiainen (1995), presented an alternative based on estimation of the bulk Richardson number, Ri_B, defined as:

$$Ri_B = \frac{gz}{\theta_V}\left(\frac{\Delta\theta + 0.61\,\theta_V \Delta q}{U_z^2}\right) \quad \text{(A33)}$$

and related as a function of the stability parameter, z/L, according to:

$$Ri_B = \frac{z}{L}\left(\frac{k\sqrt{C_{DN}}/C_{HWN} - \psi_{HW}}{\left[k/\sqrt{C_{DN}} - \psi_M\right]^2}\right) \quad \text{(A34)}$$

where it is specified that $C_{HN} = C_{WN} = C_{HWN}$. Figure A2 illustrates the relationship between the degree of atmospheric stratification (as described by both the bulk Richardson number and the Monin-Obukhov stability parameter) and the transfer coefficients scaled by their neutral value.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-257
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 20 November 2017

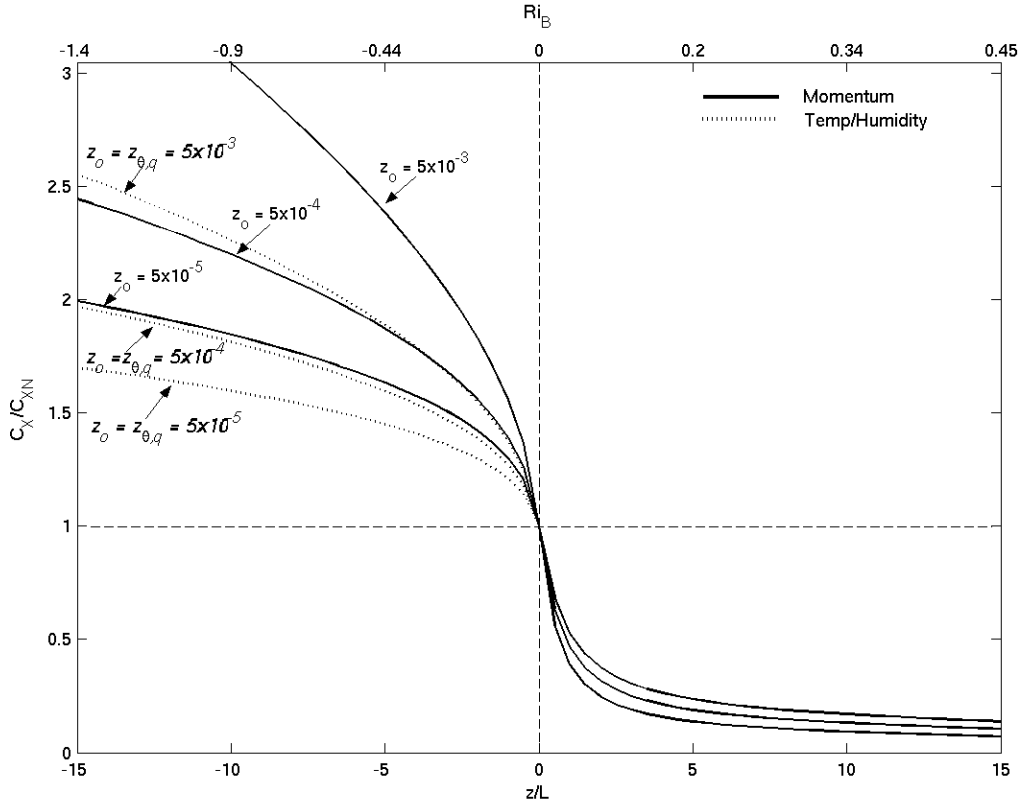

Figure A2: Relationship between atmospheric stability (bottom axis – z/L, top axis – Ri_B) and the bulk-transfer coefficients relative to their neutral value (C_X/C_{XN} where X represents D, H or W) for several roughness values (computed from Eq. A32). The solid line indicates the momentum coefficient variation (C_D/C_{DN}) and the broken line indicates humidity and temperature coefficient (C_{HW}/C_{HWN}) variation.