NEMO-GSL v1.0: An integrated hydrodynamic and ice-process model for a deep subarctic lake (Great Slave Lake, Canada)
Responses to Referee 1

November 21, 2026

We thank the referee for a thorough and critical review. The comments on the modelling context, the justification of the parameterization choices and the analysis of lake dynamics have led to substantial changes. The main changes are: (i) a new Sect. 3.1 that discusses the established lake and ocean models, their previous applications, and the reasons for and limitations of our choice of NEMO; (ii) an expanded description and justification of vertical mixing, bottom friction, light penetration and the ice configuration (Sect. 3.3); (iii) additional evaluation beyond RMSE, namely correlation, depth-resolved statistics, buoyancy-frequency profiles, agreement in stratification state and the sensitivity of the mixed layer depth (MLD) to the threshold (Sects. 4.2 and 4.4); and (iv) a new analysis of the circulation above and below the thermocline (Sects. 4.8 and 5.4, Figs. 16–17), which replaces the depth-averaged fields. In response to Referee 2, simulated conservative temperature is now converted to in-situ temperature before comparison with observations, which changed all temperature statistics. As described in our response to Referee 3, we also corrected errors in the evaporation evaluation and reran all simulations without an evaporation scaling factor. The referee's comments are in black and our responses in blue. Section, figure and table numbers refer to the revised manuscript, and all works cited here are listed in its reference list.

This manuscript documents the configuration of the NEMO model for Great Slave Lake, Canada. Great Slave Lake has the interesting feature in that its main body is relatively shallow, while a complex region of islands and channels is extremely deep (by lake standards).

While the authors provide some justification for why they are reporting on configuring a deep ocean model for a medium-large sized lake with typical depths similar to on-shelf (mean depth of 41m has been reported). NEMO seems, a priori, to be the wrong tool for the job, and its use seems to have little a priori justification.

It is especially problematic that the authors present the problem of using NEMO for lakes as a "done deal". I am aware of at least four (MIT gcm, ELCOM/AEM3D, ROMS/CROCO, FVCOM) different models in the modelling community for Canadian lakes, and my guess is that the application of these models, in terms of published papers, would outnumber those using NEMO by a factor of 10 to 1. Yet the statement made in the manuscript is that Dupont reported on applying NEMO to the Great Lakes, and then the issue is left there. I think there is a lot of rebalancing to do as far as the discussion is concerned.

We agree that the submitted text presented the choice of NEMO as self-evident, and we have rewrit-

ten this part of the manuscript. A new Sect. 3.1 ("Choice of modelling framework") reviews FVCOM, MITgcm and ELCOM/AEM3D and their applications to the Laurentian Great Lakes, Lake Geneva and Great Slave Lake itself (Leon et al., 2007). It states explicitly that any of these models could be configured for Great Slave Lake and that we do not claim NEMO is uniquely suited. The section then gives four reasons for our choice: the ice (SI^3) and passive tracer (TOP) components are developed and versioned together with the ocean core; the partial-step z-coordinate avoids the pressure-gradient errors of terrain-following coordinates over the steep slopes of the East Arm; NEMO has been evaluated in systems that share individual characteristics with Great Slave Lake, namely the Laurentian Great Lakes, the Baltic Sea and the Salish Sea; and NEMO is the lake and ocean component of Environment and Climate Change Canada's prediction systems. The section also lists the limitations of the choice: the structured grid, the sea-ice heritage of SI^3, and the hydrostatic approximation. NEMO-GSL v1.0 is presented as a documented baseline against which other frameworks can be compared.

We also agree that most of the West Arm has shelf-sea rather than deep-ocean depths. NEMO is not restricted to deep-ocean applications and is used operationally for shelf seas of comparable depth, such as the North-West European Shelf. Sect. 3.3 now gives the mean depth as 69 m for the whole lake and 41 m for the Main Basin (Evans and Muir, 2016), consistent with the value cited by the referee.

The references do a nice job on many aspects of the modelling exercise (e.g. in setting why Great Slave Lake is interesting to study, in laying out large Arctic issue), but they fall short when it comes to modelling (to be fair there is discussion of the Umlauf et al papers). This seems very strange for a methods paper on modelling!

We have added references on lake hydrodynamic, turbulence and ice modelling in Sects. 3.1 and 3.3. These cover model frameworks and their lake applications (Chen et al., 2003; Marshall et al., 1997; Hodges et al., 2000; Bennington et al., 2010; Anderson et al., 2018; Cimatoribus et al., 2018; Raju et al., 2026), coupled lake and sea-ice models (Gao et al., 2011; Losch et al., 2010; Oveisy et al., 2012; Fujisaki et al., 2012), turbulence closures and their intercomparison (Wilcox, 1988; Canuto et al., 2001; Umlauf et al., 2003; Warner et al., 2005; Reffray et al., 2015), observations of mixing in large lakes (Wüest and Lorke, 2003; Cannon et al., 2019, 2021), and bulk surface fluxes over large lakes (Charusombat et al., 2018).

The authors state that parameterizations are vital for the sharp thermoclines in likes, but then provide little detail of why NEMO's suite of parameterizations are the way to go. Why is a k-omega model appropriate? Why is something like the KPP scheme not appropriate when the thermal response is so surface trapped (figures 6 and 7). There are relevant papers discussed (The Umlauf et al papers) and what is around lines 205-210 is good. But I would want to know more on both the choice of model and how sensitive modelling is to the precise values chosen (i.e. the background values of eddy viscosity strike me as so low they are irrelevant; most dissipation will be numerical dissipation.

We have expanded the justification of the vertical mixing scheme in Sect. 3.3. KPP is not readily available in NEMO. Many NEMO configurations use the one-equation TKE scheme, but we chose the Generic Length Scale (GLS) framework (Umlauf and Burchard, 2003), a two-equation approach that contains the k–kl (Mellor–Yamada), k–ε, k–ω and generic closures within one framework and has been evaluated in stratified coastal and lake settings (Umlauf and Burchard, 2005; Warner et al., 2005; Reffray et al., 2015).

All four GLS closures were tested. Table R1 compares the k–ω closure used in NEMO-GSL with the widely used k–ε closure, using the same parameter set. k–ε gives a slightly lower RMSE for the temperature profiles, while k–ω gives a lower RMSE for MLD and LSWT, a smaller evaporation deficit and better freeze-up timing. The parameter set was calibrated with k–ω, and k–ε would need recalibration

to reach equivalent ice errors. As stated in Sect. 3.3, the differences between the two closures are small, consistent with Warner et al. (2005) and Reffray et al. (2015), who found that differences between two-equation closures are small compared with the influence of the stability functions and boundary conditions.

Table R1. Model performance with the k–ω and k–ε closures, using the same parameter set. The tests were carried out without the evaporation scaling factor (see our response to Referee 3). ME is simulated minus observed.

	k–ω		k–ε	
Variable	RMSE	ME	RMSE	ME
Temperature profiles (°C)	2.01	−0.01	1.85	+0.12
LSWT (°C)	1.66	−0.39	1.68	−0.54
MLD, $\Delta T = 0.5$ °C (m)	5.42	+0.16	5.56	+1.10
$z(N^2_{\max})$ (m)	6.63	−0.99	6.41	+0.32
Evaporation bias at IWI (%)	−16.5		−18.8	
Ice break-up error (days)	1.75		2	
Ice freeze-up error (days)	1		1.75	

Regarding the background values, we agree that the chosen values (2×10^{-6} and 2×10^{-7} m^2 s^{-1}) are small, and this is intended. As stated in Sect. 3.3, they are close to molecular, consistent with observations of near-molecular hypolimnetic mixing in large lakes (Wüest and Lorke, 2003; Cannon et al., 2021) and with the background diffusivity found to best preserve the thermocline structure in Lake Michigan–Huron (Raju et al., 2026). Vertical mixing is therefore left to the turbulence computed by GLS. Simulations with larger background values produced a more diffuse thermocline. We agree that numerical mixing may exceed the background values. As noted in Sect. 4.2, numerical diffusion is a possible contributor to the more diffuse simulated thermocline, although the model uses a fourth-order flux-corrected transport scheme for tracer advection.

What is the cutoff for this model as far as horizontal resolution? The 1 km seems like a good start, but it's not clear to me that all dynamic phenomena that is relevant to modeling is a priori hydrostatic. And what about vertical resolution during the winter when the dynamics is presumably quite sensitive?

The 1 km resolution is a compromise between computational efficiency, the sparse observations available for the bathymetry, and the resolution of lake processes (Sect. 3.3). For the summer stratification, a two-layer estimate gives an internal Rossby radius of about 1.6–2.4 km (upper layer of 8–12 m at 14–16 °C, lower layer of 25–40 m at 5–6 °C, $f = 1.28 \times 10^{-4}$ s^{-1}), consistent with the few kilometres stated in Sect. 5.4. The 1 km grid therefore resolves the basin-scale circulation of the West Arm, which is much wider than the Rossby radius, but resolves internal Kelvin waves only marginally, as now stated in Sect. 3.3. We agree that not all relevant processes are hydrostatic. As stated in Sects. 3.1 and 3.3, non-hydrostatic processes such as internal solitary waves and convective plumes are not represented explicitly, and convection is parameterized through the GLS closure and enhanced vertical diffusion. In winter, the level thickness of 1 m at the surface, increasing to 2.5 m at 40 m depth, is sufficient to represent the inverse stratification with several levels, but the thin boundary layer at the ice interface is not resolved (Sect. 3.3). As described in Sect. 4.2, the simulated winter inverse stratification is smoother and warmer at depth than observed (Fig. 8).

I was also left wondering about the overall dynamics of the lake. How does transport within the shallow main portion compare to systematic exchange with the complex, deep arms? Surely the point

of a modelling exercise like this one should, at least in part, have something to say about the actual dynamics in the lake (as opposed the statistics of error). I question reporting depth averaged circulation, given that the thermal stratification reported earlier is quite shallow.

We agree that the depth-averaged circulation obscured the shallow stratification, and we have removed it. The circulation is now analysed separately above and below the daily thermocline in each water column (Sect. 3.5). The thermocline is defined as the depth of maximum vertical temperature gradient below 3 m, and columns with a maximum gradient below 0.2 °C m^{-1} are treated as unstratified. Results are shown for 1998 and 2001, the two years with the most contrasting winds (Figs. 16–17, Sect. 4.8), and discussed in Sect. 5.4.

Three features answer the referee's question on the main basin and the deep arms:

- The main basin is dominated by a wind-modulated counterclockwise coastal circulation, including a northward coastal current that carries Slave River water along the eastern shore of the West Arm.
- Below the thermocline, flow is weak (generally < 0.01 m s^{-1}) and concentrated along slopes and constrictions. Within the East Arm it is below 0.005 m s^{-1}.
- Exchange between the arms occurs through the connecting channel and was two-layered in 1998: warm surface water moved east through the northern passages, while colder subsurface water returned west.

A full analysis of transport pathways and residence times between the basins is beyond the scope of this model description paper. As noted in Sect. 5.4, it is the subject of a companion study using passive tracers (in preparation).

What is the effect of using a thermodynamic only ice model? There seems to be a fair amount of detail (reported around line 220) in the model, but does that detail matter if the ice does not move? How sensitive is the model to the parameters?

Without ice dynamics, ice in NEMO-GSL v1.0 grows and melts in place: wind-driven drift, the opening of leads and the piling of ice against downwind shores are not represented. This is now stated in Sect. 3.3, together with the reasons for the choice. We expect the effect to be largest in the deep central basin, which freezes over later than the margins and where wind-driven ice drift and break-up have been observed (Kang et al., 2012), and it may bias break-up timing there. Once consolidated, however, the ice cover behaves largely as landfast ice, and the simulated freeze-up and break-up dates agree with satellite observations within mean absolute errors of 1 and 1.75 days, respectively. This suggests that ice timing at the lake scale is controlled mainly by thermodynamics.

The thermodynamic detail therefore matters even though the ice does not move. The multi-layer ice and snow thermodynamics, the snow thermal conductivity and the albedos determine the rates of ice growth and melt, and hence the freeze-up and break-up dates, the length of the open-water season and the heat exchange between the lake and the atmosphere. The simulated ice and snow mass is also used in the ice-load correction of the lake level for the outflow calculation (Sect. 3.4).

We did not carry out a systematic sensitivity analysis of the ice parameters, which is beyond the scope of this study. The snow thermal conductivity and the albedos of melting snow and ice were adjusted manually during calibration and proved important for reproducing the observed freeze-up and break-up timing. Including ice dynamics is identified in the manuscript as a logical extension of the configuration that could reduce the remaining break-up errors (Sects. 3.3 and 5.2).

What is known about the accuracy of the forcing product? Are there weather stations one can ground

truth against? I get that you have to use reanalysis, but a bit more information would be useful on how good one expects it to be.

We agree that this information is useful. CaSR v3.2 is built on the same reanalysis system as its predecessor RDRS (Sect. 3.4). Gasset et al. (2021) evaluated that system over 2011–2017 and found its skill to be stable over time and comparable to that of Environment and Climate Change Canada's operational forecasts, with the assimilation of land-surface data improving temperature and humidity in the lower troposphere and surface layer. No evaluation of the reanalysis over Great Slave Lake is available for our study period. Over-lake measurements at the Inner Whaleback Island site (Blanken et al., 2008) were made available to us only for 1998–1999, and the three shoreline stations (Yellowknife, Hay River and Fort Resolution) are influenced by land–lake contrasts that a 10 km reanalysis cannot resolve, so a station comparison would mainly measure that representativeness error. Instead, the forcing sensitivity experiments (Sect. 4.7) show how strongly the simulated MLD, LSWT and evaporation depend on the forcing, and the accuracy of the over-lake forcing, particularly wind speed, is discussed as a source of uncertainty in Sect. 5.5.

A different category of criticism is the over-reliance on RMSE. Of course RMSE is how one starts to measure error, but surely one cannot just stop there. Take Figure 6 for example; there are clear spatiotemporal differences that, at least to the eye, seem far more important than an overall RMSE value. I think this needs some discussion.

We agree, and we have added several analyses that go beyond a single RMSE value:

- The Pearson correlation and depth-resolved RMSE, ME and correlation for the temperature profiles (Fig. 6, Sect. 4.2). Errors peak between 5 and 15 m with little bias, indicating differences in the timing and vertical position of the thermocline rather than a systematic offset.
- A discussion of the spatiotemporal differences visible in the isotherms (Fig. 6 of the submitted manuscript, now Fig. 7). These include the earlier spring warming of the simulated hypolimnion and the timing and magnitude of storm-driven mixing events (Sect. 4.2).
- Median profiles of buoyancy frequency, which compare the vertical structure of stratification (Fig. 11, Sect. 4.4).
- The agreement in stratification state, the timing of stratification onset (Fig. 12), and the sensitivity of MLD to the threshold (Table 3, Sect. 4.4).

RMSE and ME are retained because they allow comparison with published evaluations of large-lake models.

In summary, this is the start of something good. A lot of work was done, and the model will clearly be good for something. But the modelling context needs to be improved, and a subset of the above critiques need to be addressed (the authors can't do everything, and I assume they will make judicious choices governed by what is doable).

We thank the referee for this assessment. Following the referee's advice, we have prioritised the modelling context and model choice, the analysis of lake dynamics, and the evaluation beyond RMSE.

Small points:

The manuscript does an admirable job describing the steps taken to produce a bathymetry, but the result (Fig. 4) is separated from the discussion by a large chunk of the article.

The interpolated bathymetry (Fig. 4 of the submitted manuscript) is now shown in the bathymetry methods (Sect. 3.2) as Fig. 3, directly after the input observations (Fig. 2). The interpolation error assessment remains in Sect. 4.1 (Fig. 5).

I question the need for an acronym for Great Slave Lake. GSL is generic sounding and the location is unique and deserves to be written out in full.

Agreed. "Great Slave Lake" is now written in full throughout the text and figure captions. The abbreviation is kept only in the configuration name NEMO-GSL.

The wind rose in Fig 14 is pretty small at present. Could you fill the lower right space with the wind rose and put the legend in the space above?

In the revised circulation figures (Figs. 13 and 14 of the submitted manuscript, now Figs. 16 and 17), the wind rose has been enlarged and placed in the lower-right corner of panel (b), and the legend has been moved to panel (a), as suggested. The captions now state that wind directions follow the meteorological convention, that is, the direction from which the wind blows.

The axis and axis label fonts in general are faint and small. I would increase size and perhaps make them bold.

Axis labels and tick labels have been enlarged and set in bold.