Sensitivity of MITgcm to different model parameters in application to Cayuga Lake

A. Dorostkar & L. Boegman
Department of Civil Engineering, Queen's University, Kingston, ON, Canada

P.J. Diamessis
Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA

A. Pollard
Department of Mechanical Engineering, Queen's University, Kingston, ON, Canada

ABSTRACT: The MITgcm (MIT general circulation model) has been applied to study geophysical flows with a broad range of scales. However, in the majority of applications, the accuracy of model results remains uncertain as the model has not been validated against field observations. In this paper, temperature measurements from Cayuga Lake are used to validate the MITgcm and study the sensitivity of the model to different model parameters. Both quantitative and qualitative methods are used to assess model skill. It is found that the linear equation of state (EOS) yields poorer results in comparison to polynomial EOS formulations. The vertical stratification shows a strong sensitivity to the background vertical viscosity when the KPP mixing scheme is employed. It is shown that the root mean square error between modeled and observed data is 1.9 $^{\circ}$C minimized and the observed vertical stratification is reproduced well with a Smagorinsky scheme for the horizontal eddy viscosity and a background vertical eddy viscosity of 10^{-3} m^2 s^{-1}, with a polynomial EOS.

1 INTRODUCTION

The work presented in this paper is part of a study to simulate the degeneration of internal waves, from an initial basin-scale seiche to high-frequency nonlinear internal waves (NLIWs), in a medium-sized lake. To fulfill this objective, a parallel three-dimensional (3D) code with non-hydrostatic capability, the MITgcm (Marshall et al. 1997a) is used. This model has been applied to simulations of geophysical flows over a broad range of scales, from large-scale global circulation (Adcroft et al. 2004, Marshall et al. 1997b, Nycander et al. 2007, Weijer 2005) to small-scale processes such as convection and internal waves. At global scales, due to the large domain and coarse grid, the hydrostatic solver is employed. Also, owing to lack of field observations, the model parameters such as sub-grid scale (SGS) closures and EOS formulas are typically set without performing any sensitivity analysis and calibration based on existing field data. The parameters are set in some applications to small scale processes (Marshall et al. 1997b, Vlasenko & Stachchuk 2007) in which the non-hydrostatic solver is employed. Thus, the accuracy of the various parameterizations has not been fully tested. The MITgcm has been tested through comparison with analytical solutions (Legg and Huijts 2006, Legg and Adcroft 2003) or other numerical model results (Berntsen et al. 2008), in some small scale process applications. However, the validation of model results against field observations is limited to some recent qualitative comparisons (Legg and Klymak 2008). These graphical comparisons do not provide any statistical information to judge skill assessment of the model. Thus, there is a great need for the evaluation of the different SGS models employed in the MITgcm.

In this study, a sensitivity analysis and calibration are carried out in the application of the MITgcm to simulate the basin-scale (hydrostatic) response of Cayuga Lake to wind forcing. Understanding model sensitivity is a vital initial step, thereby enabling the exploration of efficient high-resolution non-hydrostatic simulations, which will demand significant computational effort. The validated model will be used to assess the ability of the MITgcm to reproduce the

spectrum of internal waves at field scale in future studies. We thank S. Schweitzer & E. Cowen for providing field observations. The project was funded by NSERC and Queen's University.

2 METHODS

2.1 *Study site and field measurement*

Cayuga Lake (Fig. 1) is the second-largest of the glacial Finger Lakes of central New York State; it has a length of 61.4 km, maximum depth of 132 m and average width of 2.8 km. The deepest point is near the geographical center of the basin and the bathymetry is dominated by shallow mudflats at the north and more southern steeper slopes (see Fig. 1).

A series of field experiments were conducted during mid-September and mid-November of 2006 (Schweitzer et al., pers. comm.) to study the internal wave climate in Cayuga Lake. Four thermistor strings (S1, S2, S3 & S4) equipped with SeaBird SBE-39 temperature loggers were deployed (Fig. 1a). The thermistors were spaced at various depths starting from 10 m up to 80 m depth in the water column (Fig. 1b). Water temperature data were collected at 25 sec intervals. A shore based meteorological station was located near the mid-lake position, which recorded wind speed and direction at 10 min intervals at 6 m height above the water surface.

2.2 *MITgcm*

The MITgcm is a non-hydrostatic z-coordinate finite volume model that solves the incompressible Navier-Stokes equations with the Boussinesq approximation on an Arakawa-C grid. The horizontal subgrid-scale mixing is parameterized with a constant eddy viscosity or nonlinear Smagorinsky viscosities.

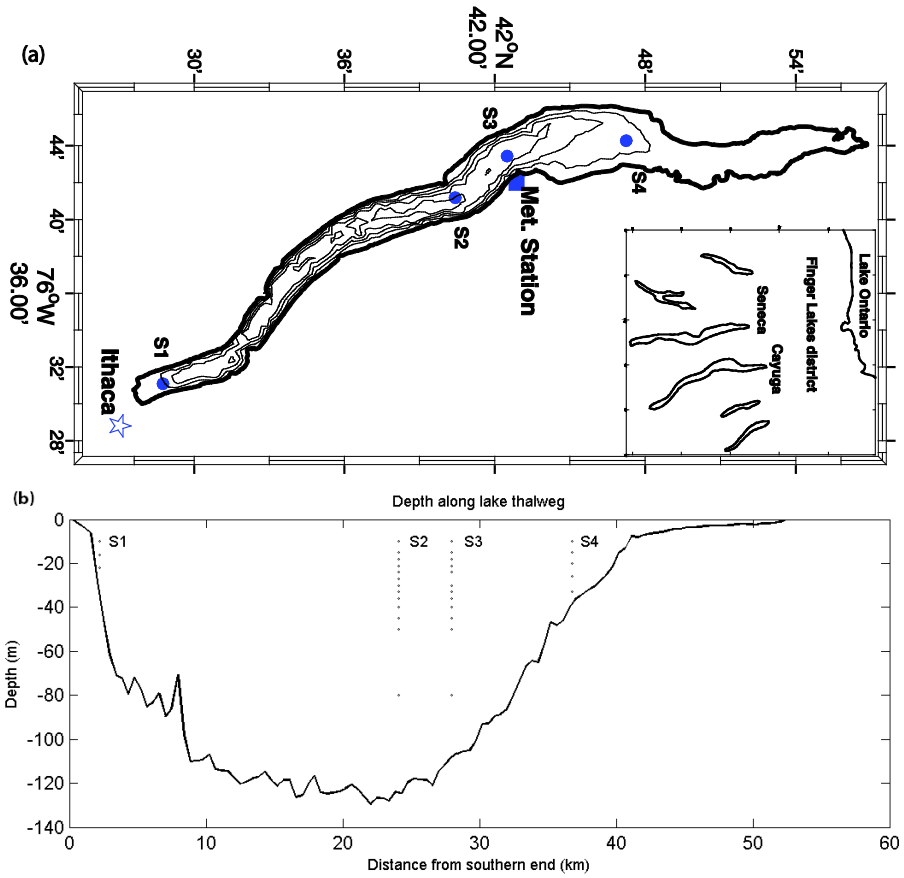

Figure 1. (a) Cayuga Lake bathymetry with depth contours (30 m interval) and location of meteorological station and temperature logger strings; (b) Cayuga Lake profile, showing the location of thermistors

The vertical sub-grid-scale mixing of scalars can be defined using a constant eddy diffusivity or a K profile parameterization (KPP) scheme. The tracer advection scheme is upwind biased third order direct space time (DST) with flux limiters. The model includes linear and polynomial equation of state (EOS) formulas. The linear EOS assumes the density only depends on temperature where the thermal expansion coefficient of water is $2x10^{-4}$ K^{-1}. The polynomial EOS formula is the (McDougall et al. 2003) equation (hereafter; MDJWF).

2.3 *Model configuration for Cayuga Lake*

All simulations are performed for 11 days, from Julian day 269 until 280 during 2006, on 10 quad-core 2.52 Ghz Sparc64 VII processors with 2 Tbytes of memory available using a time step of 40 s, a horizontal grid resolution of 450x450 m and a fixed non-uniform vertical resolution. Vertical layers are spaced every 0.5 m between 10 and 30 m depths to resolve the thermal stratification throughout the metalimnion. Near the surface and within the portion of the water column spanning depths of 30 to 50 m, a grid spacing of 1 m is used. Within the remaining portion of the water column a grid spacing of 2.95 m is employed. The model is initialized with a flat free surface, zero velocities and a horizontally uniform water temperature profile throughout the lake, taken from a 48 hr (Julian day 267 to 269) average of temperature data at S2. No-flux boundary conditions are applied for the momentum and scalar. Boundary conditions for the momentum are free-slip at the lateral boundaries, no-slip with quadratic bottom friction at the bottom and wind stress at the surface. Wind stress fluxes in zonal and meridional directions are calculated using the bulk formulae. The explicit horizontal diffusivity is set to 10^{-7} m^2 s^{-1} as the tracer advection scheme is unconditionally stable. The vertical eddy viscosities and diffusivities are computed by the KPP scheme. The other model parameters are chosen during calibration.

2.4 *Comparison methods*

Both statistical quantities and graphical comparisons are used to assess model skill. In order to obtain a quantitative picture of how well different models agree with temperature observations, Taylor diagrams (Taylor 2001) are used. This concise diagram (see Fig. 2) enables visualization of three statistical measures including linear correlation coefficient (R), standard deviation (σ), and root-mean-square error (RMSE). The angular coordinate indicates the correlation coefficient between the model and observed data. The radial coordinate is proportional to the standard deviations. The observed data lies on the horizontal axis with a correlation coefficient of unity. The linear distance from this point is equal to the unbiased RMSE (RMSE′).The term unbiased means any information about the potential bias - the difference between the means of the two fields – is removed. A relatively accurate model would be positioned near the dotted arc (having correct σ) and close to the observed field illustrating a high correlation and small RMSE′. The R and RMSE′ for a temperature time-series at a single observation point are determined by Equations 1 and 2. Additionally, the spatiotemporal RMSE is calculated for a temperature field of M=31 spatially distributed thermistors and for a period of N=2160 time steps as Equation 3:

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}(m_i-\bar{m})(o_i-\bar{o})}{\sigma_m \sigma_o} \tag{1}$$

$$RMSE' = \left(\frac{1}{N}\sum_{i=1}^{N}[(m_i-\bar{m})-(o_i-\bar{o})]^2\right)^{0.5} \tag{2}$$

$$RMSE = \left(\frac{1}{M}\sum_{j=1}^{M}\frac{1}{N}\sum_{i=1}^{N}\left(m_{ij}-o_{ij}\right)^2\right)^{0.5} \tag{3}$$

where the letter m and o indicate the model and observed temperature field, respectively, the overbar indicates the average, σ is the standard deviation.

In order to provide an overview of how models perform over time and to identify unusual data that may contribute to poor skill metrics, the time evolution of temperature isotherms obtained from different numerical experiments are also visually compared with thermistor data.

2.5 *Presentation and discussion of results*

Various sensitivity runs were carried out to determine how the model results are affected with different SGS closures and EOS formulas. A subset of these runs is shown in Table 1. Figure 2 includes Taylor diagrams for temperature time-series computed over the 10 day analysis period (JD 270 - 280) near the surface, thermocline and bottom regions of S3 and S4 for various test cases shown in Table 1. Figure 3 displays the graphical comparisons of temperature contours of Run 5, Run 6, Run 9 and field data. As illustrated in Figures 3c and 3d, the general agreement between observed and simulated data is quite satisfactory, achieving $R > 0.7$ and the spatiotemporal RMSE of 1.9 $^{\circ}$C. The model (Run 9) reproduces the baroclinic oscillations associated with the horizontal mode one internal seiche (with a period of 80 hr), upwelling of the thermocline along the shore at S1 and S4 and the formation of a progressive nonlinear surge at S2. These dynamics are characteristic of long, narrow lakes (Boegman et al. 2005) and details on their simulation will be reported elsewhere.

Table 1. List of sensitivity runs and their parameter settings

Run	Mixing closure	EOS	*ViscA$_h$ (m^2s^{-1})	*ViscA$_z$ ($m^2s^{-1)}$	*ViscC
1	Constant	Linear	1	1E-3	N/A
2	Constant	MDJWF	1	1E-3	N/A
3	Constant	MDJWF	1E-5	1E-3	N/A
4	Constant	MDJWF	1E-5	1E-5	N/A
5	Smagorinsky	Linear	Variable	1E-3	1
6	Smagorinsky	MDJWF	Variable	1E-5	1
7	Smagorinsky	MDJWF	Variable	1E-3	1
8	Smagorisnky	MDJWF	Variable	1E-3	2
9	Smagorinsky	MDJWF	Variable	1E-3	0.7

* ViscAh and VisAz are background horizontal and vertical eddy viscosities, respectively. ViscC is Smagorinsky scaling coefficient.

We first compare the results of numerical experiments which differ only in EOS (Runs 1&2 vs. Runs 5&7) (Fig. 2). The linear models (Runs 1&5) yield poorer mean correlation (R=0.5 versus R=0.7) with field data in comparison to the polynomial models (Runs 2&7). Linear models also underestimate the standard deviation of simulated water temperature by 50% on average while polynomial models are more accurate in predicting the amplitude of internal waves due to temperature fluctuations (see S2&S3 in Fig. 3a,d). As shown, Run 5 with linear EOS does not reproduce observed upwelling at S1 and S4 and internal surge formation at S2 and S3. The spatiotemporal RMSE computed with linear models is 2.3°C which is 20% larger than the RMSE for polynomial models. A number of applications of MITgcm (Marshall et al. 1997b, Nycander et al. 2007, Weijer 2005) have employed the linear EOS for computational efficiency, yet the results have not been validated against field/laboratory data. As with any linear estimate, the results are only valid for small changes in state variables and according to our findings, the accuracy of these results may be improved using a polynomial EOS.

The horizontal eddy viscosity is either constant or determined using a Smagorinsky parameterization. With the constant, changing the eddy viscosity from 10^{-5} m^2s^{-1} (Run 3) to 1 m^2s^{-1} (Run 2) has no effect on the results. As it is presented in Figure 2, Run 2 and Run 3 are coincident on the Taylor diagrams. This is consistent with the typical application of a constant horizontal eddy viscosity of 1 m^2 s^{-1} in geophysical modeling (Rueda 2003). A temperature contour comparison of results using the constant eddy viscosity (Run 2) and Smagorinsky (Run 7) shows general agreement between them (not shown here). In numerical simulations of stratified flows, the horizontal eddy viscosity does not play a very crucial role owing to the dominant effect of advection in the horizontal relative to the vertical dimension. However, the Smagorinsky model performs slightly better (~2%) in overall in terms of statistical criteria (Fig. 2). It is found that the model is not very sensitive to the dimensionless scaling coefficient (ViscC) over the range 0.1-0.7 (not shown). However, larger values of VicC (e.g. Run 8) produce large RMSE in the thermocline region, by damping out the large-scale motions. With the KPP scheme, the model is not sensitive to the background viscosity in the range of 10^{-7} to 10^{-4} m^2s^{-1}. However, the background

viscosity of 10^{-3} m^2s^{-1} in Run 3/Run 7 improves the performance of the simulation such that the spatiotemporal RMSE is reduced by 20% over Run 4/Run 6, which uses a value of 10^{-5} m^2s^{-1}, yielding a greater goodness-of-fit (Fig. 2&3). Similar to our result, no sensitivity is found to the vertical background viscosity of 10^{-6} m^2s^{-1} and 10^{-5} m^2s^{-1} in the study of temperature structure in Lake Kinneret (Pan et al. 2002). However, they did not test the effect of larger values.

3 CONCLUSIONS

Run 9 best captured the dominant features of the internal wave field in Cayuga Lake. The MITgcm was sensitive to the background vertical viscosity and EOS but not the horizontal eddy viscosity. Future work will extend the simulations to capture the non-hydrostatic internal wave field in the lake.

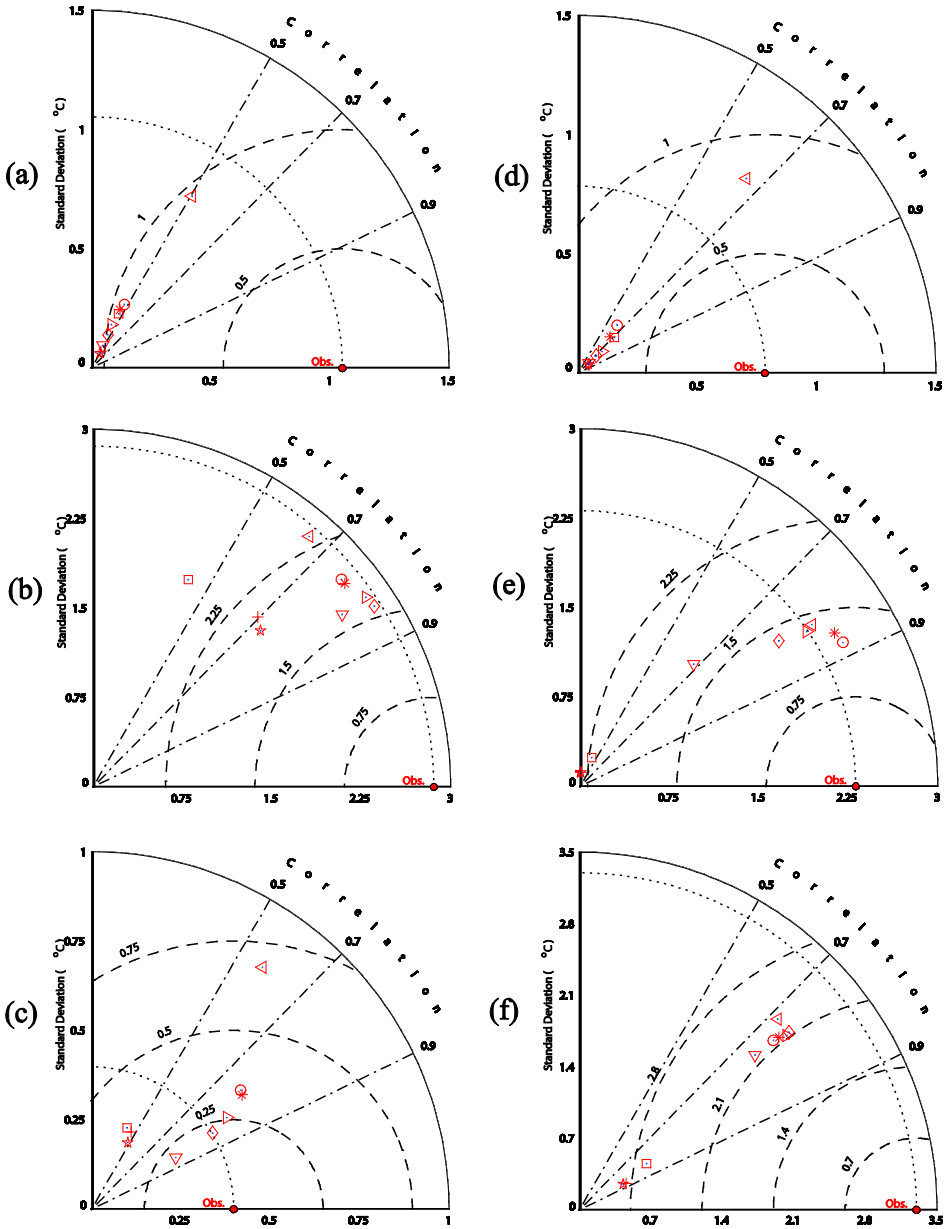

Figure 2. Taylor diagrams comparing results of various models including Run1 (+), Run2 (∗), Run3 (o), Run4 (left-pointing triangle), Run5 (pentagram), Run6 (□), Run7 (◊), Run8 (∇), Run9 (right-pointing triangle) and temperature observations for S3 at depth 10m (a), 21m (b) and 50m (c); for S2 at depth 10m (d), 15m (e) and depth 33m (d).

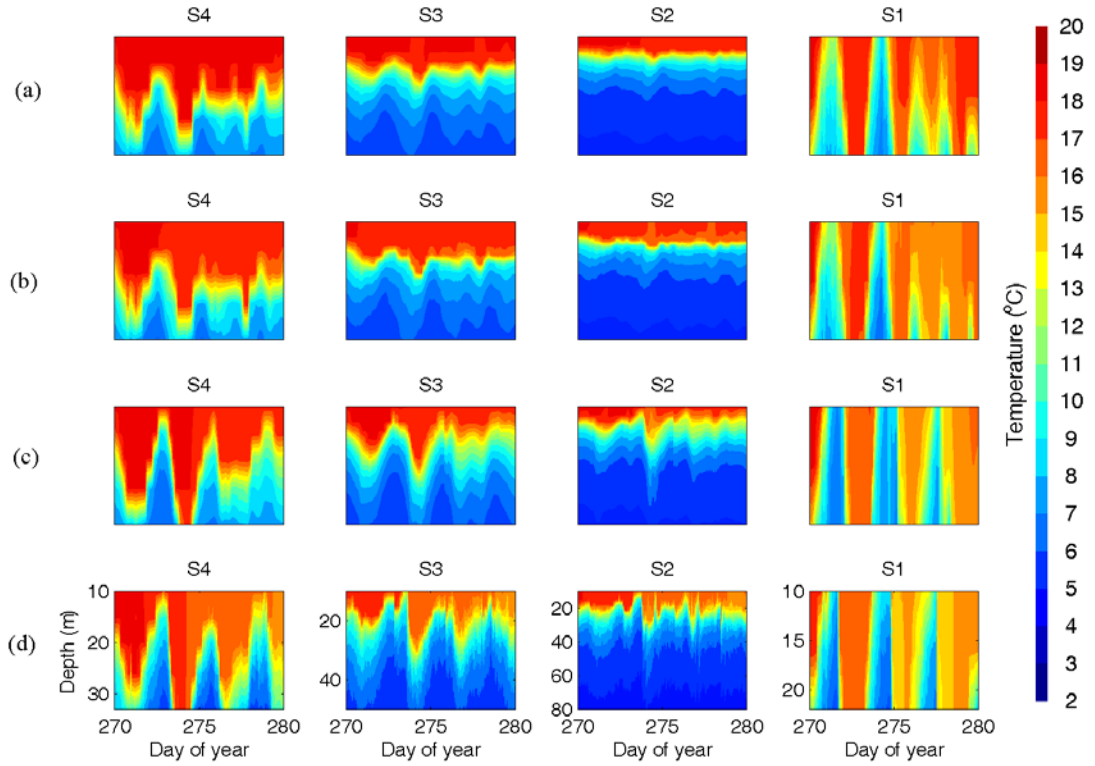

Figure 3. Cayuga Lake isotherms at S1, S2, S3 and S4 for a) Run 5, b) Run 6, c) Run 9, d) field data.

REFERENCES

Adcroft, A., Campin, J. M., Hill, C., Marshall, J. (2004). "Implementation of an atmosphere–ocean General Circulation Model on the Expanded Spherical Cube." *Mon. Weather Rev.*, 132(12), 2845-2863.

Berntsen, J., Xing, J., Davies, A. M. (2008). "Numerical Studies of Internal Waves at a Sill: Sensitivity to Horizontal Grid Size and Subgrid Scale Closure." *Cont. Shelf Res.*, 28(10-11), 1376-1393.

Boegman, L., Ivey, G. N., Imberger, J. (2005). "The Energetics of Large-Scale Internal Wave Degeneration in Lakes." *J. Fluid Mech.*, 531, 159-180.

Legg, S., and Klymak, J. (2008). "Internal Hydraulic Jumps and Overturning Generated by Tidal Flow Over a Tall Steep Ridge." *J. Phys. Oceanogr.*, 38(9), 1949-1964.

Legg, S., and Adcroft, A. (2003). "Internal Wave Breaking at Concave and Convex Continental Slopes." *J. Phys. Oceanogr.*, 33(11), 2224-2246.

Legg, S. (., and Huijts, K. M. H. (2006). "Preliminary Simulations of Internal Waves and Mixing Generated by Finite Amplitude Tidal Flow Over Isolated Topography." *Deep-Sea Res.*, 53(1-2), 140-156.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C. (1997a). "A Finite-Volume, Incompressible Navier Stokes Model for Studies of the Ocean on Parallel Computers." *J.Geophys.Res*, 5753-5766.

Marshall, J., Hill, C., Perelman, L., Adcroft, A. (1997b). "Hydrostatic, Quasi-Hydrostatic, and Nonhydrostatic Ocean Modeling." *J.Geophys.Res*, 102(C3), 5733-5752.

McDougall, T. J., Jackett, D. R., Wright, D. G., Feistel, R. (2003). "Accurate and Computationally Efficient Algorithms for Potential Temperature and Density of Seawater." *J. Atmos. Ocean. Technol.*, 20(5), 730-741.

Nycander, J., Nilsson, J., Döös, K., Broström, G. (2007). "Thermodynamic Analysis of Ocean Circulation." *J. Phys. Oceanogr.*, 37(8), 2038-2052.

Pan, H., Avissar, R., Haidvogel, D. B. (2002). "Summer Circulation and Temperature Structure of Lake Kinneret." *J. Phys. Oceanogr.*, 32(1), 295-313.

Rueda, F. J. (2003). "Dynamics of Large Polymictic Lake. II: Numerical Simulations." *J. Hydraul. Eng.*, 129, 92.

Taylor, K. E. (2001). "Summarizing Multiple Aspects of Model Performance in a Single Diagram." J.Geophys.Res, 106(D7), 7183-7192.

Vlasenko, V., and Stashchuk, N. (2007). "Three-Dimensional Shoaling of Large-Amplitude Internal Waves." J.Geophys.Res, 112, 1–15.

Weijer, W. (2005). "High-Frequency Wind Forcing of a Channel Model of the ACC: Normal Mode Excitation." *Ocean Modelling*, 9(1), 31-50.