Appendix B. Cayuga Lake Model (CLM-2D) Setup and Calibration for Cayuga Lake

The information presented in this Appendix is a summary of previously published reports: (1) Phase I: Monitoring and Modeling Support for a Phosphorus/Eutrophication Model for Cayuga Lake (288p; UFI 2014) and (2) Phase 2 Final Report: A Phosphorus/Eutrophication Water Quality Model for Cayuga Lake, New York (227p; UFI 2017). For more information, please see the aforementioned reports.

B1. Introduction

This Appendix describes the setup, calibration, and performance of the Cayuga Lake Model (CLM-2D), a hydrodynamic and water quality model developed for Cayuga Lake, New York. CLM-2D is a hybrid model developed using the CE-QUAL-W2 hydrothermal transport model (Cole and Wells 2015) and water quality sub-models developed for Cayuga Lake by the Upstate Freshwater Institute, Syracuse, New York.

B2. Description of CLM-2D

Sub-models of CLM-2D

A model is a theoretical construct that assigns numerical values to parameters and relates external inputs or forcing conditions to system variable responses (Thomann and Mueller 1987, Chapra 1997). CLM-2D is a two-dimensional model composed of a hydrothermal/transport model and water quality sub-models. The two-dimensional hydrothermal/transport sub-model used in CLM-2D is the hydrothermal/transport sub-model in CE-QUAL-W2, a public access model developed by the U.S. Army Corps of Engineers (Cole and Wells 2015). The hydrothermal/transport sub-model was setup and tested during Phase 1 (version 3.70; UFI 2014; Gelda et al. 2015) and subsequently upgraded to version 3.72 for the final CLM-2D model. The two-dimensional transport model simulates the thermal stratification regime and mixing/transport processes in the vertical and longitudinal dimensions for Cayuga Lake. The hydrothermal/transport sub-model was calibrated using 2013 observations and validated with observations from the 1998-2012 period (UFI 2014; Gelda et al. 2015).

Hydrothermal/Transport Sub-model

CE-QUAL-W2 is a public domain two-dimensional (longitudinal and vertical) hydrodynamic and water quality model (Cole and Wells 2015, http://www.cee.pdx.edu/w2). The model assumes lateral homogeneity within a segment of a waterbody and is therefore ideally suited for long and narrow waterbodies such as rivers or narrow lakes (Cole and Wells 2015). CE-QUAL-W2 is capable of predicting water surface elevations, velocities, temperature, and several water quality constituents. The model represents a waterbody using multiple longitudinal segments and multiple vertical layers within each segment. Resolution for Cayuga Lake was 25 longitudinal segments (~ 2,450m) and 1 m vertical layers from the water surface to the lake bottom (UFI 2017).

Nutrient-Phytoplankton Water Quality Model (CLM-2D)

The inflow concentrations for the water quality model follow the same formatting and daily input frequency of CE-QUAL-W2. However, the model structure and state variables used in CLM-2D differ from those used in CE-QUAL-W2. The water quality model is described in Section B5. CLM-2D includes sub-models representing algae and Chlorophyll-a, zooplankton, the effects of dreissenid mussels, and four major algal constituents: (1) carbon (C), (2) phosphorus (P), (3) nitrogen (N), and (4) silica (Si). Sub-models are also included for dissolved oxygen, minerogenic particles and optics (e.g., Secchi depth). For a more detailed description of this model please see UFI 2017.

B3. Overview of Model Setup and Data Requirements

Overview of Model Setup and Data Requirements

The general data requirements of the hydrothermal/transport model are: (1) geometric data (lake bathymetry, model cell dimensions, elevation, area, volume); (2) meteorological data (air and dew point temperature, wind velocity and direction, cloud cover or solar radiation); (3) hydrologic data (tributary inflows, outflows, and water surface elevation); (4) nutrient concentrations and temperatures of the lake and its tributaries; (5) hydrodynamic and kinetic coefficients; and (6) other data such as water withdrawals or discharges (Table B1). CE-QUAL-W2 uses laterally

averaged two-dimensional (vertical and longitudinal) equations of fluid motion (Edinger and Buchak 1975). Inherent to this framework is the assumption of uniform lateral mixing in the cross-channel direction. The basic equations that describe water movement and the movement of materials (such as nutrients) are described in detail in UFI 2017.

The primary drivers for CLM-2D fall into one of three types: (1) meteorological, (2) hydrologic, and (3) constituent loading. Meteorological measurements are critical to drive the hydrothermal/transport sub-model and incident light is utilized in the phytoplankton growth sub-model. These measurements are available from a proximate location on Cornell campus (hourly since 1987), and from a site on the lake at its southern end (15 min. intervals) since 2011 (Table B2). Several of the major tributaries that enter the lake as well as lake water surface elevation are presently continuously gaged by the United States Geological Survey (USGS; Table B3). Estimates of overall tributary inflow and lake level are embedded in the hydrologic budget maintained within the model. Descriptions of constituent loading are presented in Section B3.

Table B1. Data needs for CE-QUAL-W2 and CLM-2D lake modeling.

Data Requirement	Data Type	Purpose
1	Bathymetric map of lake – a three dimensional map of lake length, width, and depth	Define dimensions of model segments and layers
2	Hourly meteorological records (air temperature, dew point temperature, wind speed, wind direction, solar radiation, and cloud cover)	Define meteorological forcings
3	Time series of inflow flow rates, water temperatures, and concentrations of water quality constituents for all inflows (tributaries, direct drainage, point sources, etc.)	Define boundary conditions
	Time series of outflow flow rates and locations of all outflows (outlets, withdrawals, etc.)	Define downstream boundary conditions.
	Water surface elevation records	Model calibration
4	In-lake water temperature and water quality records	Model calibration
5	Measured kinetic or estimated model coefficients from field data (if available)	Defining initial parameter values

Geometric Data/Model Bathymetry

CLM-2D requires the same bathymetric data as CE-QUAL-W2 (Cole and Wells 2015). Bathymetric data define the physical size and shape of a lake and consists of a number of vertical layers and longitudinal segments. The grid formed by these layers and segments (cells) is called the computational grid. The geometry of the computational grid is determined by: (1) longitudinal spacing, (2) vertical spacing, and (3) average cross-sectional width. Segment boundaries were first established on contour maps for the lake. Dimensions for each of the computational cells were then obtained from analysis of the bathymetric data.
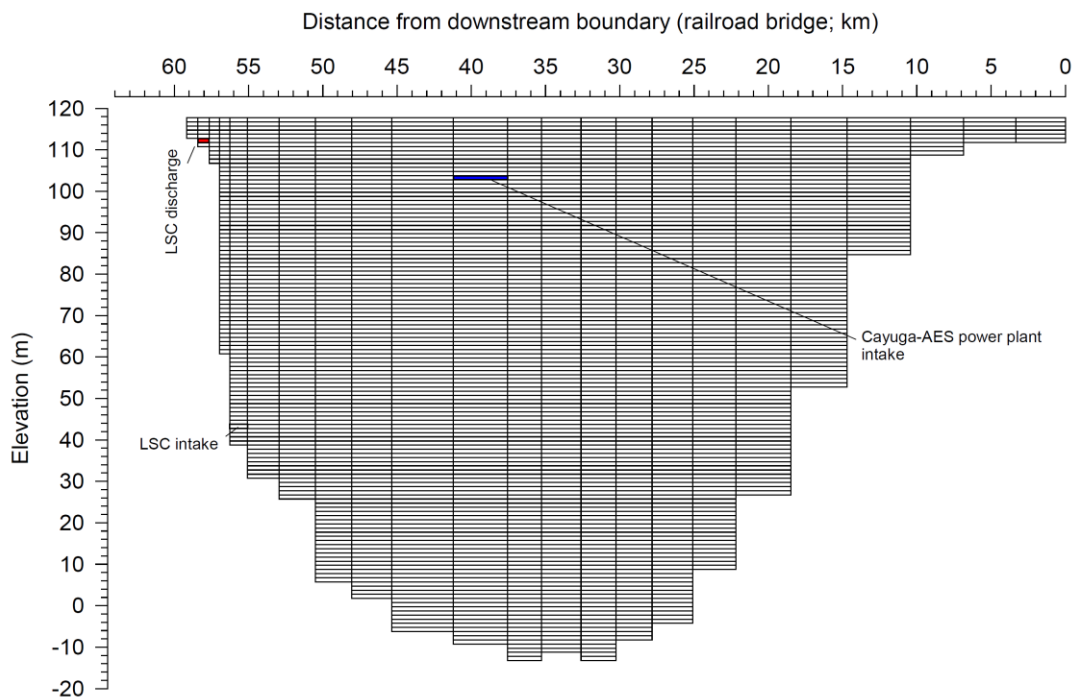
Meteorological Data

CLM-2D requires hourly average air temperature, dew point temperature, wind velocity and cloud cover (or solar radiation) data to calculate surface heat exchange and wind stress. Three meteorological stations are located near Cayuga Lake; two stations belong to Cornell University (the "pile cluster" meteorological station and the Game Farm Road meteorological station) and the third station is a NOAA station that collects data from the Ithaca Airport. These data are available for different periods (Table B2).

(a)

Figure B1a (from Figure 7-2, UFI 2017). Plan view of the CLM-2D model geometry with model segment number, calibration year monitoring site, and point sources. Inset is the southern shelf and LSC water intake. WWTP = wastewater treatment plant (or facility), IA – Ithaca Area, CH = Cayuga Heights, LSC = Lake Source Cooling

(b)

Figure B1b (from Figure 7-3, UFI 2017). Side view of the CLM-2D model geometry with the location of the LSC intake, LSC discharge, and Cayuga-AES power plant intake. Figure is orientated from south (left) to north (right).

Table B2. Meteorological station summary.

Station Name	Latitude	Longitude	Availability	Elevation (ft)	Notes
Piling cluster	42.46	-76.52	10/27/2011-12/31/2013	380	10 minute frequency; missing data (Tair and Tdew 1/3/2013 – 5/13/2013 filled in from Ithaca Airport
Game Farm Road	42.44	-76.45	1987-2013	950	Hourly frequency; missing data (0.8% days) were filled in from Ithaca Airport data
Ithaca Airport	42.49	-76.46		1080	

Flow Budget

CE-QUAL-W2 (Cole and Wells 2015) requires specification of daily average inflows from tributaries, outflows, withdrawals, and water surface elevation. A hydrologic flow budget was constructed for Cayuga Lake for the period 1987 – 2013 from the available inflow and lake volume data. Imbalances in the hydrologic budget were attributed to uncertainty in the estimation of ungauged inflows and outflows as well as potential inflow from the Seneca River to the north end of Cayuga Lake. A summary of the tributaries monitored in this study, ranked according to watershed area, is presented in Table B3. Detailed explanation of the flow budget, including procedures to estimate flow in ungagged watersheds can be found in UFI 2017.

Table B3. Tributary watershed areas and 2013 mean flow rate.

Tributary	USGS Gage No.	Watershed Area (acres)	Percent of Total Watershed (%)	2013 Mean Flow (m³/s)
Fall Creek	04234000	81,792	17.7	5.95
Cayuga Inlet Creek	04233255	59,528	12.9	2.69
Salmon Creek	0423401815	57,773	12.5	3.75
Taughannock Creek	-a	42,749	9.3	3.11
Sixmile Creek	04233300	33,137	7.2	2.09
Ungaged tributaries	-a	187,355	40.5	11.98
Total	-	462,260	100	29.56

[a] estimated from product of Fall Creek flow and Taughannock Creek to Fall Creek watershed areas
[b] estimated from product of gaged flow and ratio of total watershed area to gaged watershed area

Briefly, the overall flow budget is shown in the equations below:

$$Qin - Qout, total = \Delta$$

$$(Qin, g + Qin, ung + Qin, pt) - (Qout, pt + Qout) = \Delta$$

where: Qin (total flow into the lake) is the sum of gaged stream flow (Qin, g), ungaged stream flow (Qin, ung), and the point source inflows (Qin, pt.). Total flow out of the lake is the sum of all point source withdrawals (Qout, pt) and outflow from the lake (Qout), Δ = the change in water volume in the lake which is estimated from water surface elevation and bathymetry data.

Fall Creek, the largest tributary to Cayuga Lake, has a watershed area of 330.9 km², which represents approximately 17.7 percent of the total Cayuga Lake watershed area (Figure B2; Haith et al. 2012). In addition to the four gaged tributaries (Table B3), Taughannock Creek was also monitored for constituents in Phase 1 (UFI 2014) but was ungaged for flow. Taughannock Creek flows were estimated using Fall Creek flows and the ratio between the Fall Creek and Taughannock Creek watershed areas. Point source inflows (Qin, pt) and outflows (Qout,pt) were measured or estimated. The change in water volume (Δ) was estimated from a seven-day average of the daily measured USGS water surface elevation and known bathymetry. The total ungaged inflow (Qin, ung) was estimated as the product of the gaged inflow times the ratio of the ungaged watershed area (from Haith et al., 2012) to the gaged watershed areas (Taughannock Creek estimates included as gauged).

A flow budget was used to solve for outflows from the lake Qout to the Seneca River. In 2013, the flow budget predicted negative outflows from the lake approximately 14% of the time.
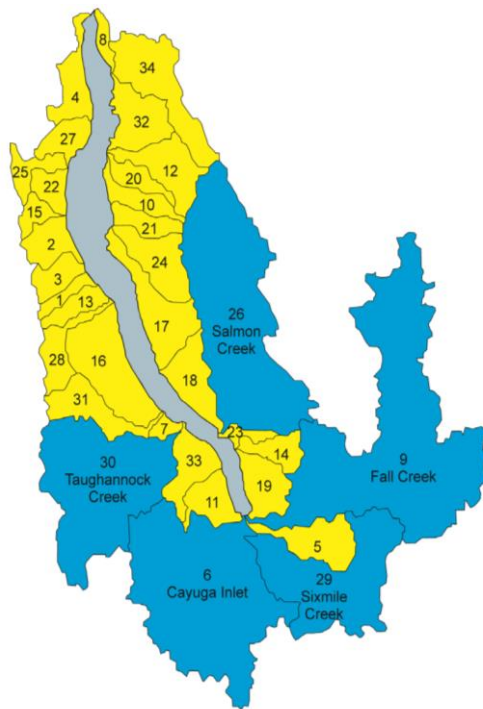
Figure B2. Revised map from Haith et al. (2012; Figure 2) with gaged watersheds colored blue, ungaged watersheds colored yellow. Watershed number and related model segments are shown in Table B4.

An assumption was made that allowed the negative outflows to be set to the value of the Seneca River inflow to the lake. The Seneca River can flow into the north end of Cayuga Lake and the USGS verified that does occur during certain times of the year, typically Fall and Spring when the elevation of the downstream lock (Mud Lock) is adjusted (W. Coon, personal communication). This assumption was verified by conducting a separate flow budget for 2015. In 2015, flows were measured both upstream of the Seneca River entering Cayuga Lake and downstream of Cayuga Lake's outlet. The difference between the two gages in 2015 was compared to the estimated Seneca River inflow in 2015 as calculated by the flow budget. For the April – October interval of 2015 the estimated outflows tracked the measured outflows well, corroborating the use of this flow budget technique. Observed and predicted water surface elevations matched well for the CLM (Figure B3).

Figure B2 is a modified figure from Haith et al. (2012) showing the five gaged tributaries colored in blue and the remaining 29 ungaged tributaries in yellow. For the CLM, these 29 areas were aggregated to 15 ungaged sub-basins (Table B4).

Table B4. The listing of the ungaged watersheds and the model segment that these tributaries enter the into the lake.

Ungaged Tributary Number	CLM-2D Segment No.	Watershed Name (from Haith et al. 2012)
ug1	2	Cascadilla C.
ug2	7	Glenwood C. area; Lansing area
ug3	8	Gulf C. area
ug4	9	Willow C.; Minnegar C.
ug5	11	Lake Ridge Point area; Trumansburg C.; Cayuga View area
ug6	14	King Ferry Sta. area
ug7	15	Sheldrake C.; Interlaken area
ug8	16	Grovers/Powel Creek area

Ungaged Tributary Number	CLM-2D Segment No.	Watershed Name (from Haith et al. 2012)
ug9	17	Barnum Creek area
ug10	18	Bloomer/Mack Ck. area
ug11	19	Little C. Area; Paines C.; Hicks Gully; Big Hollow area
ug12	20	Glen/Dean Ck. Area; Red C.; McDuffie Town area
ug13	21	Great Gully; Lavanna area; Schuyler C. area; Union Springs area
ug14	22	Canoga C. area
ug15	23	Cayuga Village area; Yawger C.
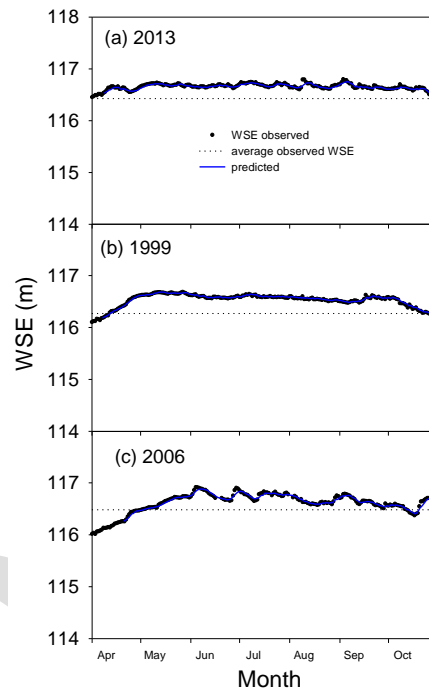
Figure B3. Time series of predicted (CLM -2D) and observed water surface elevations in Cayuga Lake in (a) 2013, (b) 1999, and (c) 2006.

Inflow Temperatures

CLM-2D requires daily inputs of stream temperature. During Phase 1 (UFI 2014), the daily stream temperatures used for the distributed inflows were assumed to be the same as those measured at a USGS site near Cayuga Lake. During Phase 2, the hydrothermal/transport model was updated to use estimates of daily temperatures based on in-stream measurements for 2013 and estimates for other years based on site-specific relationships. As part of the Phase 1 work, UFI routinely monitored stream temperatures on the five main tributaries during 2013. UFI had previously developed a method of estimating daily stream temperatures from daily air temperatures and a routine set of monitoring data (UFI 2001, UFI 2007). This method was used to estimate daily stream temperatures in 2013 for input to the hydrothermal/transport model. A final assumption that the ungaged tributaries had the same temperature as Salmon Creek was based on measurements at several of the ungaged streams made on a single day in June 2013.

Stream temperatures were not routinely measured for the validation years. UFI developed an air temperature - stream temperature regression for each of the streams monitored in 2013. These regressions and the measured air temperatures were used to develop stream temperatures for the other hydrothermal/transport model needs (UFI 2017)

Water Quality Loadings

Daily time series of water quality constituent (nutrients and sediment) concentrations are needed for tributary inflows in CLM-2D. Concentrations of were derived from tributary measurements or estimates (UFI 2014). Table B5 is a summary of required water quality inputs and methods applied to derive inflow concentrations for CLM-2D state variables.

Table B5. Inflow concentrations to CLM-2D with descriptions

Inflow Concentration	Description	Notes/Derivation
SRP	Soluble reactive phosphorus	from tributary measurements and calculations, see UFI 2014
NH4	Ammonium nitrogen	from tributary measurements and calculations, see UFI 2014
NO3	Nitrate nitrogen	from tributary measurements and calculations, see UFI 2014
DSi	Dissolved reactive silica	from tributary measurements and calculations, see UFI 2014
PSi	Particulate silica	estimated to be 0, see UFI 2014
LDOM	Labile Dissolved Organic Matter	Fraction of the from sum of Particulate and Dissolved Organic Carbon, from tributary measurements, see UFI 2014, 2017
RDOM	Refractory Dissolved Organic Matter	
LPOM	Labile Particulate Organic Matter	
RPOM	Refractory Particulate Organic Matter	
DO	Dissolved oxygen	see UFI 2014, 2017
LDOM_P	Labile Dissolved Organic Phosphorus	from tributary measurements and calculations, particle analysis, and bioavailability assays, see UFI 2014, 2017
RDOM_P	Refractory Dissolved Organic Phosphorus	
LPOM_P	Labile Particulate Organic Phosphorus	
RPOM_P	Refractory Particulate Organic Phosphorus	
LDOM_N	Labile Dissolved Organic Nitrogen	see UFI 2014, 2017
RDOM_N	Refractory Dissolved Organic Phosphorus	
LPOM_N	Labile Particulate Organic Phosphorus	
RPOM_N	Refractory Particulate Organic Nitrogen	
LPIP	Labile Particulate Inorganic Phosphorus	from tributary measurements and calculations, particle analysis, and bioavailability assays, see UFI 2014, 2017
RPIP	Refractory Particulate Inorganic Phosphorus	
PAVm1	Projected (Particle) Area per unit Volume – size class 1 ()	used to determine lake clarity and sediment transport, from individual particle analysis, tributary measurements and calculations, see UFI 2014, 2017
PAVm2	Projected (Particle) Area per unit Volume – size class 2 ()	
PAVm3	Projected (Particle) Area per unit Volume – size class 3 ()	
PAVm4	Projected (Particle) Area per unit Volume – size class 4 ()	

Atmospheric Loadings

Due to the size of the Cayuga Lake watershed and the importance of external nutrient loading and internal nutrient recycling, atmospheric inputs to Cayuga Lake were considered negligible and not modeled as part of the CLM-2D.

Water Quality Constituents

The in-lake water quality constituents modeled in the CLM-2D can be found in Section B5.

Bottom Sediments

Cayuga Lake's water column is fully oxygenated and does not experience anoxic conditions. Therefore, sediment nutrient release from these conditions was considered to be zero and not modeled in the CLM-2D.

Sediment Temperature

Sediment temperatures were set in the control file under variable TSED to be 10°C. This value was set as part of the hydrothermal model calibration (Gelda et al., 2015).

Initial Values

The model was initialized by the measurements made on the first day of sampling (April 8) in 2013. The setup and testing and data needs for this hydrothermal model CE-QUAL-W2 as well as final model coefficients are described in detail in the Phase 2 report (UFI, 2017). The water quality model for Cayuga Lake was CLM-2D and its parameterization is presented in Section 7.6 of the Phase 2 report (UFI, 2017). The state variables and units are listed in Table 7-6. Conceptual diagrams are presented in Figure 7-18 through 7-25. List of model drivers is presented in Table 7-12. All mass balance equations are listed in Appendix 2 of the Phase 2 Report. Table A2-1 is a full listing of all model coefficients used in the CLM-2D model calibration of Cayuga Lake including the coefficient symbol, description, value (where applicable) and unit (UFI, 2017).

Simulation Period

The simulation period for the CLM is January 1, 1998 through December 31, 2013.

B4. Calibration, Validation and Performance of the Hydrothermal Model

Section of Validation Years

In Phase 1 the hydrothermal/transport model was calibrated for 2013 and validated for 1998-2012. In Phase 2 the upgraded hydrothermal/transport model was validated for two years, 1999 and 2006. The two validation years represent a wide range of hydrologic conditions out of the 16-year study period, with 1999 ranking as the 12th wettest summer (15th on an annual basis) and 2006 ranking as the wettest summer (7th on an annual basis). Values of these metrics for the calibration year of 2013 were generally between those measured in 1999 and 2006. The wide range of meteorological forcing conditions included in the calibration and validation data sets represents a robust test of the hydrothermal/transport model.

Hydrothermal/transport Model Calibration and Validation

Model testing was based on comparisons of model predictions with measured: (1) vertical temperature profiles at multiple locations in the lake (Figure B1a), (2) signatures of oscillations in stratified layers and intrusions of hypolimnetic waters into surface layers (upwelling events) from high frequency temperature measurements in the Southern End, and (3) signatures of tributary entry.

The coefficients used for calibration and validation of the hydrothermal/transport model are shown in Table B5. These are the recommended default values except for wind sheltering (set to 1.0) and the Chezy coefficient (set to 70). Applications for numerous lakes and reservoirs under a wide variety of conditions have shown the hydrothermal/transport model generates remarkably accurate temperature predictions using default values when provided with accurate geometry and boundary conditions. Another important parameter, the light extinction coefficient, was determined from site-specific measurements of the underwater light (UFI 2017).

Table B5. Hydrothermal/transport coefficients in CE-QUAL-W2.

Coefficient	Symbol	Model Values
horizontal eddy viscosity	Ax	1 m2/sec

| horizontal eddy diffusivity	Dx	1 m2/sec
Chezy coefficient (all segments)	Ch	70 m0.5/sec
wind sheltering coefficient (all segments)	Wsc	1.0
fraction of incident solar radiation absorbed at the water surface	β	0.45
coefficient of bottom heat exchange	CBHE	0.3 W/m2/°C

Evaluation of Hydrothermal/transport Model Performance

Simulation of temperature by the hydrothermal model is a test that the model is simulating transport of heat (and therefore mass in the water quality model) in both the vertical and longitudinal directions in the lake. Temperature also regulates a number of biological processes in the lake.

The primary basis for evaluation of the hydrothermal/transport model comparisons of predictions with observations. Goodness of model fit was based on both visual inspection of model predictions to observed data and statistics, including Root Mean Square Error (RMSE). A RMSE of 1 degree C is considered sufficient model performance. Examples of the model fit are presented for site 5 (Figure B4), the primary water quality monitoring site (UFI 2014), and site 3 (Figure B5), a site with a long-term monitoring record (1998-2012). Similar plots for 2013 at the other monitoring sites are provided in UFI 2017 and in Appendices C of this TMDL. The hydrothermal model simulated observations well for the calibration year of 2013.
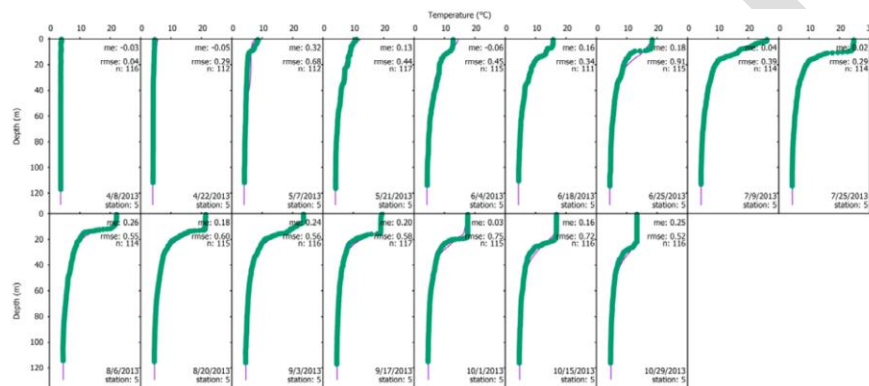
Figure B4 (from Figure 7-11, UFI 2017). Comparisons of predicted and observed 2013 temperature profiles for Cayuga Lake, site 5. Mean errors (me), root mean square errors (rmse), and number of observations (n) are included for reference. Line = model predictions, green circles = observed temperatures.
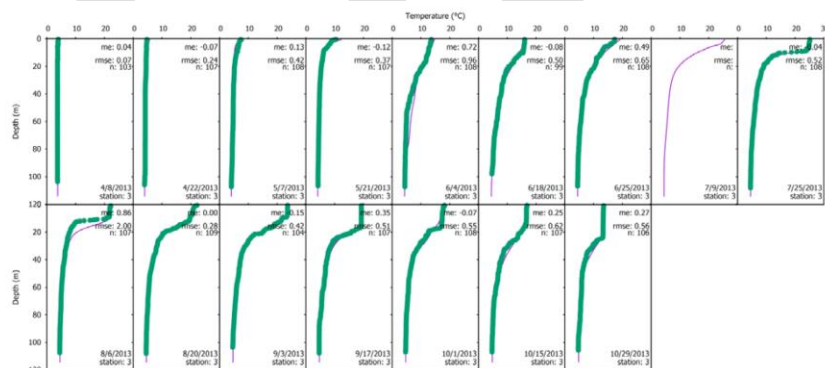
Figure B5 (from Figure 7-12, UFI 2017). Comparisons of predicted and observed 2013 temperature profiles for Cayuga Lake, site 3. Mean errors (me), root mean square errors (rmse), and number of observations (n) are included for reference

Model fits for 2013 are also presented as time series, with temperature observations and simulations shown for multiple depths at each of the nine monitoring sites (Figure B6). The model tracked the observed temperatures

sufficiently at all sites and depths. The model also performed well in simulating temperatures for the two years of model validation, 1999 and 2006 (Appendix C).

B5. Nutrient-Phytoplankton water quality model

Conceptual framework: Background and approach

The overall model utilized for the Cayuga Lake TMDL is an integrated two dimensional hydrothermal/transport and water quality model utilizing the hydrothermal/transport portion of CE-QUAL-W2 (version 3.72; Cole and Wells 2015) and a separately developed water quality model described in UFI 2017. This integrated model is referred to as CLM-2D (Cayuga Lake Model - 2D). The inflow concentrations for the water quality model followed the same formatting and daily input frequency of CE-QUAL-W2.

The constituents and characteristics predicted by the water quality model are described as the state variables (Table B6). The model also includes several derived constituents, which are calculated from the state variables (Table B7). Multiple forms of P are predicted, including particulate and dissolved fractions of nutrients, which are partitioned according to labile (subject to reactions/transformations) and refractory (not subject to reactions/transformations), and organic versus inorganic, components. Phytoplankton biomass and organic carbon (C) are simulated, with multiple forms of C (dissolved versus particulate, labile versus refractory) predicted. Chlorophyll-a (Chl-a), a surrogate of phytoplankton biomass, is derived as the product of simulated phytoplankton biomass (ALG) and the Chl-a:ALG ratio. Two groups of phytoplankton are modeled, diatoms (ALG1) and other algal taxa (ALG2). Total phosphorus (TP) was derived by summing the simulated dissolved and particulate forms of phosphorus. Secchi disk depth (ZSD) was predicted by the optics sub-model (UFI 2017).

Multiple metrics of sediment were simulated, including PAV, turbidity, and suspended solids. Optical metrics were predicted by the optics sub-model. Nitrate+nitrite (NOx) and silica (Si) were added to the Phase 1 list of model state variables because both had distinctive depletion signatures in the pelagic waters of the lake (UFI 2014). Dissolved oxygen (DO) is included as a state variable because it is needed for certain reactions. However, no effort was made to calibrate this parameter because DO is not a water quality issue for Cayuga Lake as there is no evidence of oxygen depletion in the lake's lower waters.
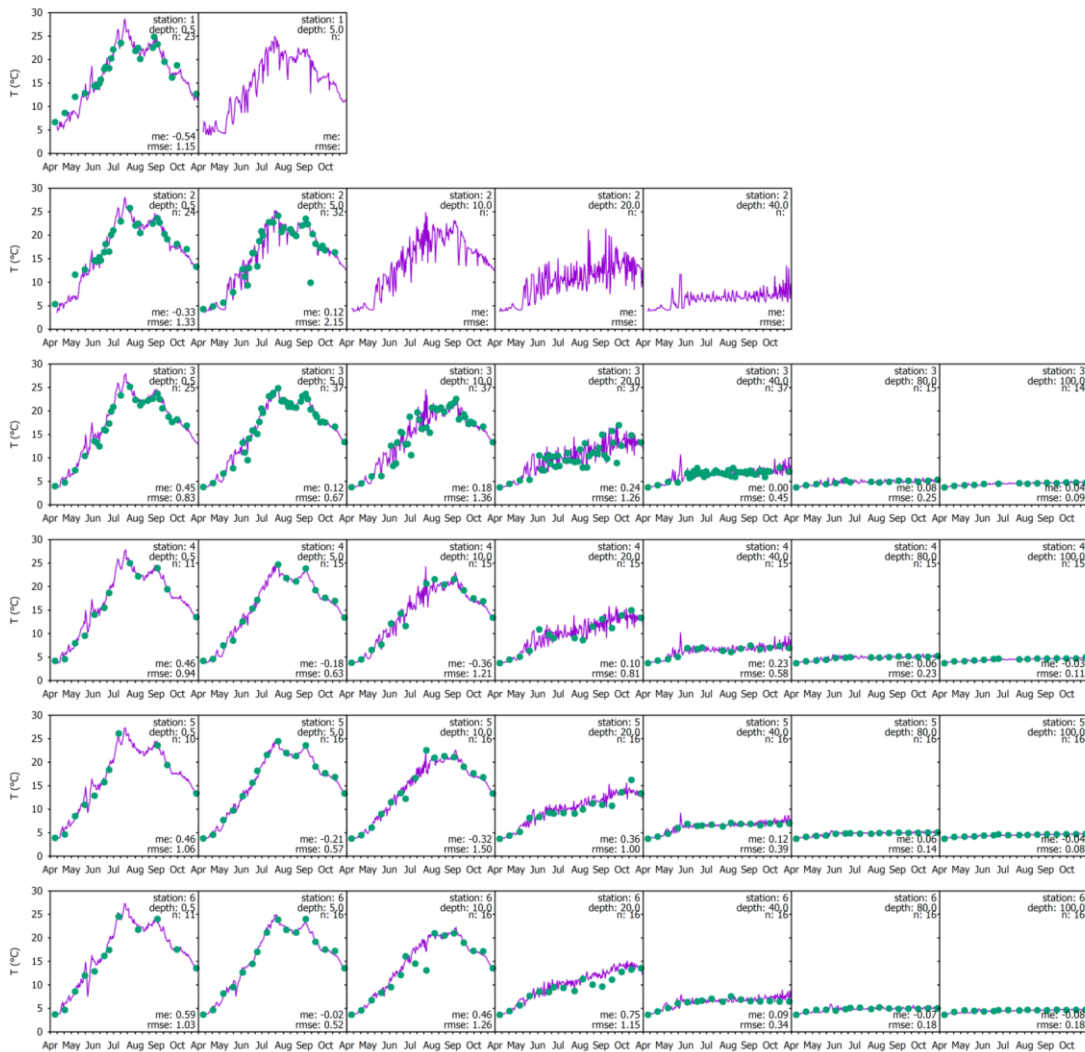
Figure B6 (from Figure 7-13, UFI 2017). Time series of predicted and observed temperatures for 2013 at nine monitoring sites and multiple depths (0.5, 5, 10, 20, 40, 80, and 100 meters – if applicable) in Cayuga Lake. Mean errors (me), root mean square errors (rmse), and number of observations (n) are included for reference. Sites 1-2 are from the Southern End segment, sites 3-6 are from the Main Lake, Mid-South segment.

Table B6. Listing of CLM-2D state variables.

Symbol	Description	Input/Output unit
T	temperature	°C
Alg1, Alg2	algae in terms of carbon	µg C/L
DO	dissolved oxygen	mg O2/L
Carbon		
LDOC	labile dissolved organic carbon	mg C/L
RDOC	refractory dissolved organic carbon	mg C/L
RPOC	labile particulate organic carbon	mg C/L
RPOC	refractory particulate organic carbon	mg C/L
CO2	carbon dioxide	mg C/L
Nitrogen		
NH3	total ammonia	µg N/L
NOX	sum of nitrate plus nitrate plus nitrite	µg N/L

Symbol	Description	Input/Output unit
LDON	labile dissolve organic nitrogen	µg N/L
RDON	refractory dissolve organic nitrogen	µg N/L
LPON	labile particulate organic nitrogen	µg N/L
Phosphorus		
SRP	soluble reactive phosphorus	µg P/L
LDOP	labile dissolve organic phosphorus	µg P/L
RDOP	refractory dissolve organic phosphorus	µg P/L
LPOP	labile particulate organic phosphorus	µg P/L
RPOP	refractory particulate organic phosphorus	µg P/L
LPIP	labile particulate inorganic phosphorus	µg P/L
RPIP	refractory particulate inorganic phosphorus	µg P/L
DSi	dissolved silica	mg Si/L
PSi	particulate silica	mg Si/L
Zooplankton		
Zoo1	zooplankton carbon, modeled or fixed	mg C/L
Mussels (fixed not modeled)		
MusDW	Mussel dry weight, fixed not modeled	g DW/m2

Table B7. Listing of CLM-2D derived variables (calculated from state variables).

Symbol	Description	Input/Output unit
Chl	chlorophyll a	µg /L
N:P	ratio of nitrogen to phosphorus	µg N/µgP
Carbon		
DOC	dissolved organic carbon	mg C/L
POC	particulate organic carbon	mg C/L
TOC	total organic carbon	mg C/L
Nitrogen		
DON	dissolved organic nitrogen	µg N/L
PN	particulate nitrogen	µg N/L
TDN	total dissolved nitrogen	µg N/L
TN	total nitrogen	µg N/L
Phosphorus		
TDP	total dissolved phosphorus	µg P/L
DOP	dissolved organic phosphorus	µg P/L
PP	particulate phosphorus	µg P/L
TP	total phosphorus	µg P/L

Carbon sub-model

Dissolved components of the carbon sub-model include carbon dioxide (CO_2), labile dissolved organic carbon (LDOC), and refractory dissolved organic carbon (RDOC). Particulate forms include zooplankton, algal carbon, labile particulate organic carbon (LPOC), and refractory particulate organic carbon (RPOC). Labile and refractory forms are differentiated by decay rates, which were determined in model calibration. Organic carbon is an important regulator of lake metabolism (Wetzel 2001, Hanson et al. 2003). CLM-2D uses POC as the primary metric of phytoplankton biomass. Sinks for algal carbon include grazing by zooplankton, mortality, settling, and ingestion by dreissenid mussels.
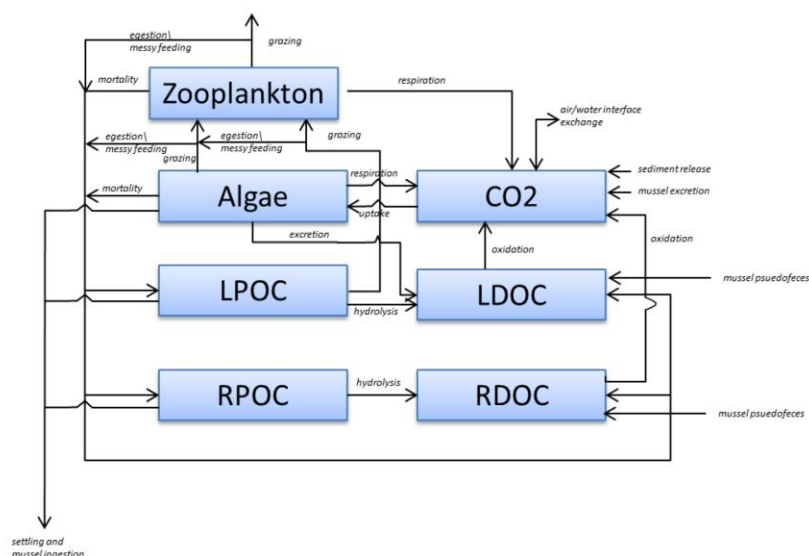
Illustration of the Carbon sub-model

Nitrogen sub-model

Dissolved forms included in the nitrogen (N) sub-model are ammonia (NH_3), nitrate+nitrite (NOx), labile dissolved organic nitrogen (LDON), and refractory dissolved organic nitrogen (RDON). The model also tracks four particulate forms of N: zooplankton, algae, labile particulate organic nitrogen (LPON), and refractory particulate organic nitrogen (RPON). Labile and refractory forms are differentiated by decay rates, which were determined in model calibration. Although both ammonia and nitrate can be used to support algal growth, ammonia is preferred for energetic reasons (Wetzel 2001). Ammonia concentrations are low in Cayuga Lake, and algal demand for N is met primarily by nitrate.

Algal N is lost through respiration (i.e., dark respiration) and excretion (i.e., photorespiration) processes to the NH_3 and LDON pools. Particulate forms of N that are lost to settling and ingestion by dreissenid mussels include algae, LPON, and RPON.
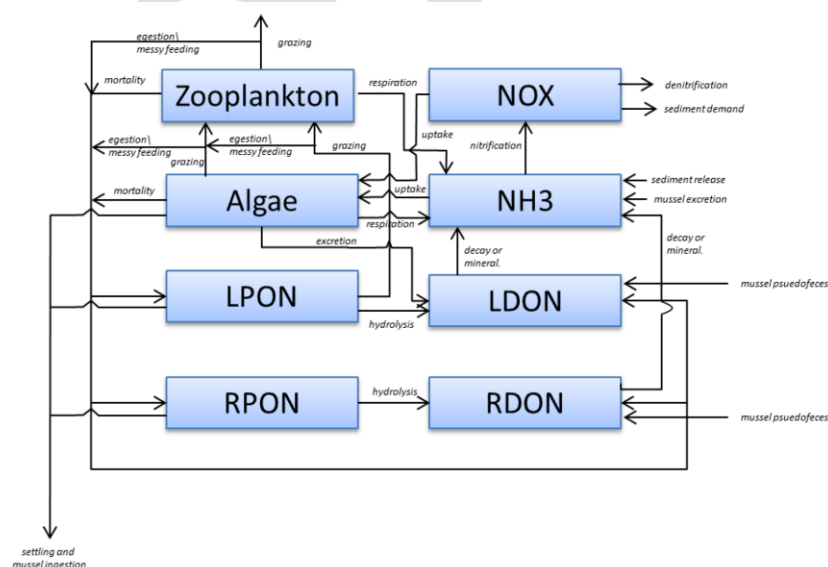

Illustration of the Nitrogen sub-model

Dissolved components of the phosphorus (P) sub-model include soluble reactive P (SRP), labile dissolved organic P (LDOP), and refractory dissolved organic P (RDOP). Particulate forms include zooplankton, algal carbon, labile particulate organic carbon (LPOC), and refractory particulate organic carbon (RPOC). Labile and refractory forms were differentiated by decay rates, which were specified according to the results of P bioavailability assays. Soluble reactive P supports algal growth, which is largely limited to the epilimnion of the lake because of limited light penetration. Sources of SRP to the water column include microbial decay of LDOP and RDOP, respiration/decay of algal phosphorus, zooplankton respiration, and dreissenid mussel excretion.

Particulate P in the form of algae and non-algal particles is lost from the water column due to settling and ingestion by dreissenid mussels. Settling velocities were determined through calibration, with lower values for algal P and higher values for non-living particulate components. The external loading of PP and DOP was partitioned according to the outcome of system-specific bioavailability experiments described in the Phase 1 report (UFI 2014). The primary modeling performance target for P was the summer average TP concentration in the upper waters, consistent with the NYSDEC guidance value for TP.
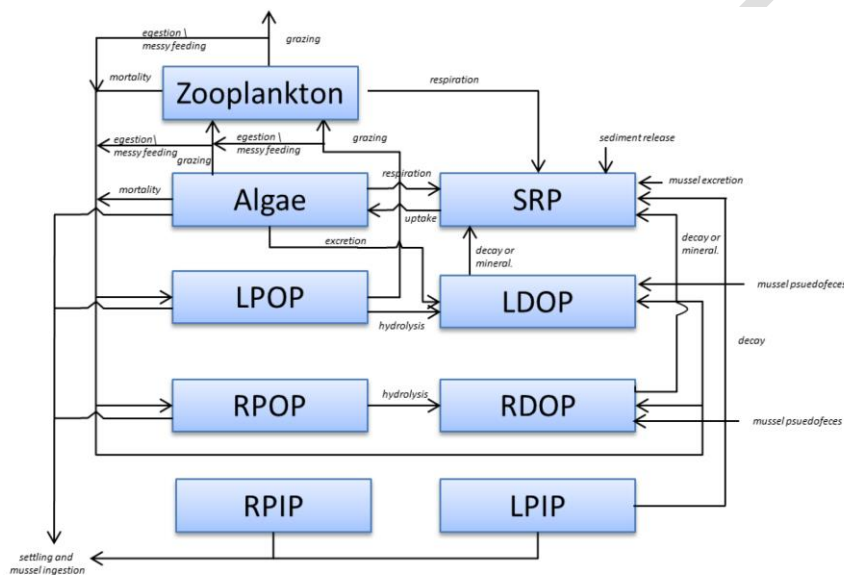

Illustration of the Phosphorus sub-model

Algal biomass in lakes in the North Temperate Zone is typically limited by a combination of phosphorus availability and seasonally intense zooplankton grazing (Wetzel 2001). However, diatoms, which use dissolved silica to form frustules, can also be limited by the availability of silica. For this reason two algal groups are modeled in CLM-2D, diatoms (ALG1) and other algae (ALG2). Algal growth is limited by temperature, light and nutrient availability.

Inorganic forms of carbon, nitrogen, and phosphorus are used to support algal growth. Although ammonia is the form of nitrogen preferred by algae, nitrate is used as an alternative when ammonia is unavailable. Sinks for algae include grazing by zooplankton, mortality, settling, and ingestion by dreissenid mussels. The processes of algal mortality and excretion transfer algal carbon to particulate and dissolved organic forms in the water column.

The primary metric of algal biomass in CLM-2D is particulate organic carbon (POC). The modeling goal for algal biomass was simulation of major seasonal dynamics and the summer average in the upper waters. The concentration of chlorophyll-a (Chl-a) is not simulated directly, but estimated as the product of the state variable POC and the Chl-a:POC ratio. Simulation of Chl-a was also a target of the initiative, at a time scale of summer average, consistent with the TMDL lake Chl-a targets. This is consistent with the known dependence of Chl-a on species composition, ambient light, and other environmental conditions (Reynolds 2006). Indeed, the Chl-a:POC ratio has been reported

to be dependent on not only light availability but nutrient status (Chalup and Laws 1990, Laws and Chalup 1990, Hecky et al. 1993).
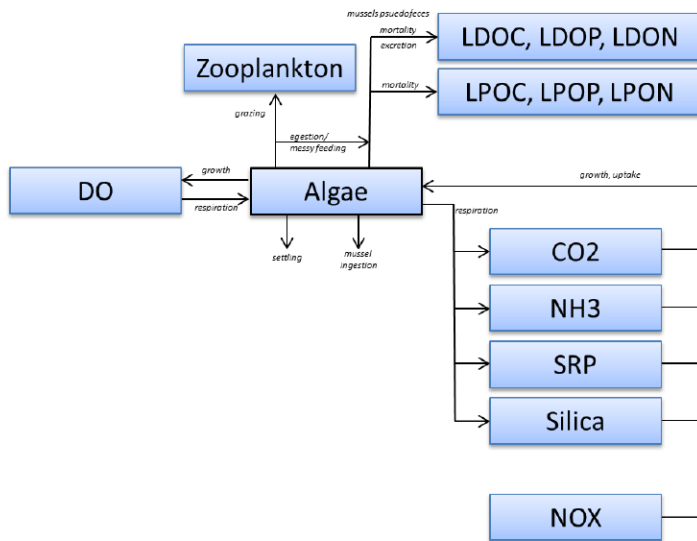
Illustration of the Algae (Chl-a) sub-model

Zooplankton sub-model

A zooplankton sub-model was included in CLM-2D to accommodate the effects of grazing on the algal community of Cayuga Lake. Zooplankton are modeled as a single group that consumes algae. Zooplankton respiration recycles algal nutrients (CO_2, NH_3, SRP) back to the water column. Labile forms of dissolved and particulate organic matter are produced as a result of zooplankton mortality.

Illustration of the Zooplankton sub-model

Modeling the effects of dreissenid mussels

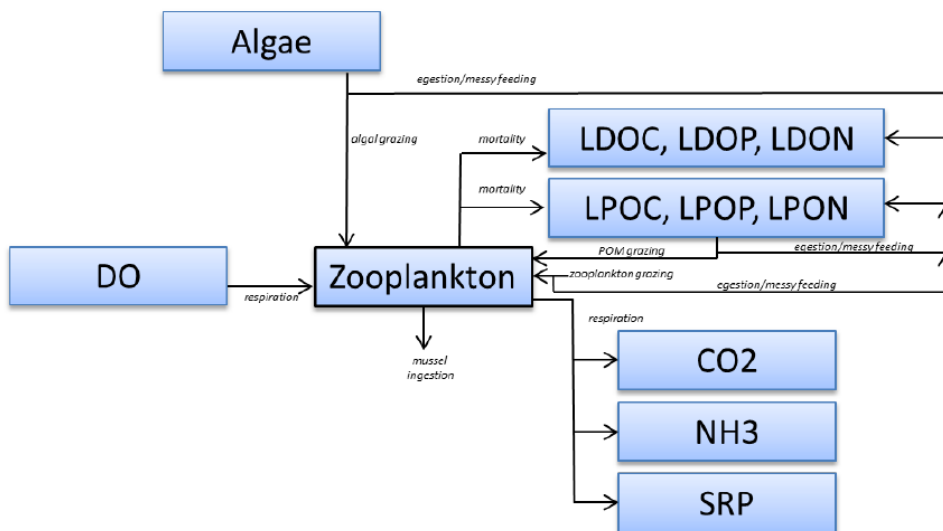The water quality sub-model was updated to accommodate the water quality impacts of dreissenid mussels found on the floor of Cayuga Lake. The model simulates the impact of dreissenid mussels on the water column by removing particulate constituents and converting a fraction of the particulates to dissolved constituents (e.g., SRP). However, the growth and mortality of the mussels were not modeled. Instead, the mussel biomass measured in 2013 as part of Phase 1 (UFI 2014, https://www.dec.ny.gov/docs/water_pdf/cltacdrei.pdf) was used as a model driver. Vertical profiles of areal density (dry weight mass per unit area of lake bottom, gDW/m^2) were developed at each sampling location for both zebra and quagga mussels. A single dreissenid mussel group was formed by summing the measured biomass of the two species. These profiles were assumed to be representative of the biomass within the model segment that the sample site was located in. The data were then spatially interpolated, both vertically and horizontally, to obtain vertically detailed profiles of dreissenid mussel density for each of the 25 model segments. The filtering rate of dreissenid mussels was determined through calibration, guided by literature reviews, laboratory experiments performed during Phase 1 (UFI 2014).
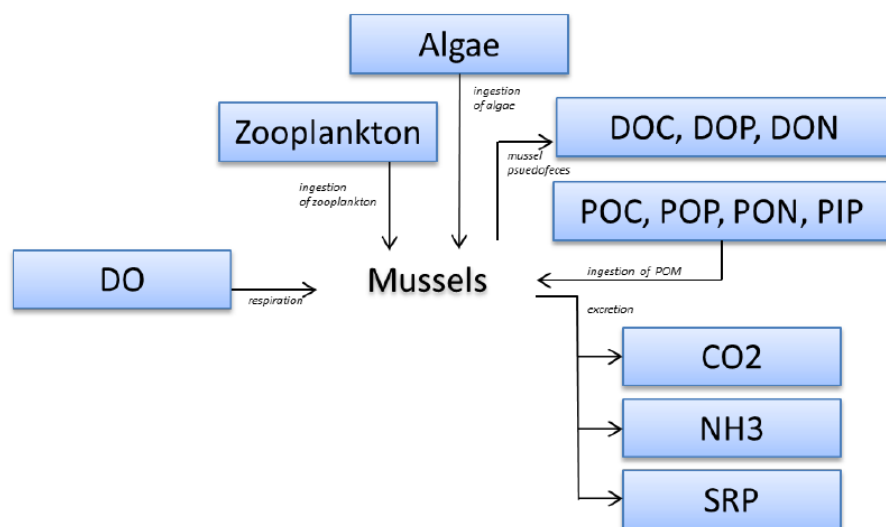

Illustration of the Dreissenid mussel sub-model

Silica sub-model (not shown)

The silica sub-model was included in CLM-2D to allow for simulation of diatoms as a separate algal group. Dissolved silica (DSi), which is used in the formation of diatom frustules, can limit diatom growth when depleted to low levels (Reynolds 2006). Sources of DSi to the water column include tributary loading and hydrolysis of particulate silica (PSi). Diatom mortality, zooplankton egestion, and messing feeding by zooplankton are the primary sources of Psi to the water column. PSi is lost to the water column through settling and ingestion by dreissenid mussels.

Dissolved oxygen sub-model (not shown)

As discussed previously, modeling of dissolved oxygen (DO) was not a priority for CLM-2D because DO is not a water quality issue for Cayuga Lake. However, the DO sub-model was operative in CLM-2D because a DO concentration is necessary for certain reaction in other sub-models. Sources of DO in the oxygen sub-model include photosynthesis (largely limited to the epilimnion) and reaeration (epilimnion, only). Oxygen sinks include algal respiration, zooplankton respiration, dreissenid mussel respiration, nitrification, and oxidation of DOC.

Minerogenic particle submodel (not shown)

As described in the Phase 1 report (UFI 2014, Section 5) and UFI publications (Gelda et al. 2016a and b, and Peng and Effler 2015), minerogenic particles delivered to Cayuga Lake from its watershed play an important role in metrics of water quality in the lake, including phosphorus, turbidity, clarity and light penetration. The key model state variable is the projected area of minerogenic particles per unit volume (PAVm).

Optics submodel (not shown)

The optics sub-model provides predictive capabilities for optical metrics of water clarity, as represented by Secchi depth (SD) and the attenuation coefficient for scalar irradiance (K0(PAR)). SD is a primary trophic state and water quality metric of concern for lacustrine systems, including Cayuga Lake. K0(PAR) is important as it specifies the light available at various depths to support photosynthesis and phytoplankton growth.

Water quality modeling protocols

The time step of hydrologic, material loading, and meteorological forcing function inputs to the water quality model is daily. The computational time step of the model calculations is one hour. The model was initialized by the measurements made at sites 1-9 on the first day of sampling (April 8) in 2013. Coefficient values were selected based on earlier work on Cayuga Lake, from the literature, or based on professional judgement and accepted limnological paradigms (UFI 2017).

Development and specification of water quality model drivers

Inflow concentrations

Constituent loads were estimated from a combination of observed concentrations (C), and those estimated from flow (Q) measurements, as described by C-Q relationships (UFI 2014). These estimates were developed using the FLUX32 software that provides flow and concentration estimates at a daily time step. Loading estimates for periods without regular tributary monitoring of concentrations were estimated from the C-Q relationship developed from the 2013 data set (UFI 2014).

Constituent concentrations for all tributary and point source inflows are a critical form of input for CLM-2D. Concentrations of various constituents were measured in 2013 at the mouths of five Cayuga Lake tributaries and in the Inlet channel, as described in the Phase 1 QAPP (UFI 2014). The methods used to calculate loads were documented in detail in the Phase 1 final report (UFI 2014, Prestigiacomo et al. 2016) and in Section 3 of UFI 2017. The resulting loads were divided by the flows to develop tributary-specific inflow concentrations.

In-lake calibration and validation data sets

The model calibration data set was collected in 2013 (UFI 2013). Data analyses and summary of findings from 2013 are documented in detail in the Phase 1 final report (UFI 2014). The validation data sets for CLM-2D rely heavily on data collected as part of Cornell University's long-term (1998-2012) monitoring program for the Lake Source Cooling facility (https://energyandsustainability.fs.cornell.edu/util/cooling/production/lsc/default.cfm). Data was collected at seven sites on the southern shelf and one deeper water site. Three of these monitoring sites corresponded to the locations of sites 1, 2, and 3 from the 2013 monitoring program. Measurements from these sites serve as the validation datasets for 1999 and 2006.

Summary of non-direct measurements

Table B8 summarizes these non-direct measurements utilized in the both phases of the modeling project.

Table B8. Non-direct measurements utilized in the Cayuga Lake Modeling Project.

No.	Data Type	Source of data	How used
1.	stream flows	United States Geological Survey (USGS)	used in flow budget and as input to the model
2.	watershed areas	Haith et al. 2012	used to adjust USGS flows and partition unmonitored flows that were utilized in a flow budget and as an input to the model
3.	meteorological data	Cornell University – pile cluster data	driver of the hydrothermal/transport model
		Cornell University – Game farm Road data	
		National Oceanic & Atmosphere Administration (NOAA)	

No.	Data Type	Source of data	How used
4.	point source flows and constituents	Cornell University – LSC based lake monitoring	flow and P data as model inputs, tNH3 as model input for WWTPs only
		IAWWTP, biweekly P data, 1995-2013, DMR data sets 2009-2013	
		CHWWTP, DMR data sets 2000-2013 for P, 2009-2013 others	
		ASE power plant	
		Minor WWTP, DMR data sets 2009-2013	
5.	stream temperatures	CSI ~2000-2013 (stream dependent)	validation of 2013 UFI air temperature vs. creek temperature regressions used to estimate creek temperatures between measurement days
6.	stream constituents	UFI 2003-2006 TP, TDP, SRP, Tn data	model inputs for years 2003-2006
		CSI ~2000-2013 (stream dependent) for TP, t-NH3, NOX, TSS, Tn	validation of 2013 concentration/flow regressions used to estimate constituents between measurement days
		DEC 2007 TP, DOC	validation of 2013 concentration/flow regressions used to estimate constituents between measurement days
		LSC based lake monitoring data	validation data sets
7.	historical limnological information – phosphorus, clarity and plankton	earlier studies by UFI and Cornell University	validation data sets utilized to develop model grid
8.	bathymetric data	Cornell University	Set-up of hydrothermal/transport model

Water Quality Model Calibration, 2013

Model performance is evaluated primarily through comparisons of model predictions with in-lake observations. Target performance thresholds for CLM-2D are provided in Table B9 for TP and Chl-a. These performance thresholds were applied on a summer average basis for lake upper waters, consistent with regulatory standards and TMDL goals.

Table B9. Targeted thresholds of model performance for multiple metrics of interest.

Predicted Metric	Targeted Thresholds of Performance[1] % Error[2]
TP	< 25%
Chl-a	< 50%

[1] summer (June-September) average values for the upper waters

[2] % Error = absolute value of (prediction – observation)/observation ×100

Observed and predicted summer average concentrations are presented in Table B10-B12 for modeled and observed TP and Chl-a concentrations for the Southern End, Main Lake, Mid-South, and Main Lake, Mid-North segments respectively.

Table B10. Comparisons of model results with performance criteria for the Southern End segment, 1998-2013.

Year	Observed TP (µg/L)	Modeled TP (µg/L)	Percent Error (%)	Observed Chl-a (µg/L)	Modeled Chl-a (µg/L)	Percent Error (%)
1998	24.2	19.5	19%	5	6	32%
1999	14.8	11.7	21%	5	4	7%
2000	19.6	19.5	1%	5	6	10%
2001	20	15.4	23%	5	5	2%
2002	20.7	20.2	2%	5	5	6%
2003	13.7	17.6	28%	7	6	15%
2004	21.8	24.2	11%	5	7	49%
2005	19	15.4	19%	5	6	16%
2006	24.9	28.9	16%	7	7	4%
2007	23.2	12.7	45%	5	5	6%
2008	17.2	13.8	20%	6	5	18%
2009	16.5	13.8	16%	5	5	4%
2010	15.7	11.5	27%	6	5	30%
2011	16.5	24	45%	5	6	19%
2012	15.3	11.4	25%	4	5	15%
2013	27	21.4	21%	4	6	32%
Average	**19.4**	**17.5**	**21%**	**5**	**6**	**16%**

Table B11. Comparisons of model results with performance criteria for the Main Lake, Mid-South segment, 1998-2013.

Year	Observed TP (µg/L)	Modeled TP (µg/L)	Percent Error (%)	Observed Chl-a (µg/L)	Modeled Chl-a (µg/L)	Percent Error (%)
1998	14.7	13.5	8%	5	5	0%
1999	9.8	8.3	15%	5	3	28%
2000	11.6	12.9	11%	5	4	8%
2001	14.1	9.5	33%	5	3	27%
2002	14.1	11.7	17%	5	4	25%
2003	10.6	10.9	3%	5	5	15%
2004	14.2	13.8	3%	5	5	2%
2005	12.6	11.7	7%	5	5	7%
2006	15.2	16.6	9%	8	6	22%
2007	13.4	10.8	19%	7	5	29%
2008	12.2	10.9	11%	7	5	33%
2009	11.6	10.7	8%	7	5	29%
2010	13	10	23%	6	4	24%
2011	14.5	15	3%	7	6	23%
2012	12.3	10.1	18%	6	5	18%
2013	15	12.1	19%	5	5	9%
Average	**13.1**	**11.8**	**13%**	**6**	**5**	**19%**

Table B12. Comparisons of model results with performance criteria for the Main Lake, Mid-North segment, 1998-2013.

Year	Observed* TP (µg/L)	Modeled TP (µg/L)	Percent Error (%)	Observed Chl-a (µg/L)	Modeled Chl-a (µg/L)	Percent Error (%)
1998	12	11	8%	6	4	34%
1999	11	6	44%	4	2	37%
2000	8	9	15%	3	3	12%
2001	10	7	24%	7	2	63%
2002	10	8	22%	5	3	45%
2003	10	9	9%	5	4	19%
2004	8	11	42%	not available	4	not available
2005	12	10	22%	3	4	23%
2006	11	12	17%	5	5	10%
2007	not available	9	not available	not available	4	not available
2008	not available	9	not available	not available	4	not available
2009	not available	9	not available	not available	4	not available
2010	not available	9	not available	not available	4	not available
2011	not available	13	not available	not available	5	not available
2012	not available	9	not available	not available	4	not available
2013	14	10	27%	3	4	47%
Average	**11**	**10**	**23%**	**4**	**4**	**32%**

*Observed Data from Makerawicz 2007.

B6. Conclusions

The Cayuga Lake Model (CLM) was used to support development of a phosphorus TMDL for Cayuga Lake. Model performance criteria were successfully met for TP and Chl-a, and other water quality parameters (UFI 2017).