REMOTE SENSING DATA ASSIMILATION IN WATER QUALITY NUMERICAL

MODELS FOR SIMULATION OF WATER COLUMN TEMPERATURE

Shuangshuang Xie

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree of
Master of Science
in the Department of Earth Sciences,
Indiana University

September 2011

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

_____

Meghna Babbar-Sebens, Ph.D., Chair

_____

Lin Li, Ph.D.

Master's Thesis
Committee

_____

Luoding Zhu, Ph.D.

DEDICATION

To the delight of my life, my forever love: my grandfather.

If you could see, you would easily tell the day and night

You could exactly grab my hand within the crowd

If you could see, I would drive you everywhere in the world

And give you a warm hug right in front

If you could see, what you loved might be totally different

I would be your eyes

To take you enjoy all the seasons

I would be your eyes

To hold your hand and walk through the crowd

I would be your eyes

To travel through the ocean of hydrology

As long as I am your eyes

You will see the lovely world in front

All in front of you

ACKNOWLEDGEMENTS

ABSTRACT

Shuangshuang Xie

REMOTE SENSING DATA ASSIMILATION IN WATER QUALITY NUMERICAL

MODELS FOR SIMULATING WATER COLUMN TEMPERATUER

Numerical models are important tools for simulating processes within complex natural systems, such as hydrodynamics and water quality processes within a water body. From decision makers' perspectives, such models also serve as useful tools for predicting the impacts of water quality problems or develop early warning systems. However, accuracy of a numerical model developed for a specific site is dependent on multiple model parameters and variables whose values are attained via calibration processes and/or expert knowledge. Real time variations in the actual aquatic system at a site necessitate continuous monitoring of the system so that model parameters and variables are regularly updated to reflect accurate conditions. Multiple sources of observations can help adjust the model better by providing benefits of individual monitoring technology within the model updating process. For example, remote sensing data provide a spatially dense dataset of model variables at the surface of a water body, while in-situ monitoring technologies can provide data at multiple depths and at more frequent time intervals than remote sensing technologies. This research aims to present an overview of an integrated modeling and data assimilation framework that combines three-dimensional numerical model with multiple sources of observations to simulate water column temperature in a eutrophic reservoir in central Indiana. A variational data assimilation approach is investigated for incorporating spatially continuous remote sensing observations and

spatially discrete in-situ observations to change initial conditions of the numerical model. This research addresses the challenge of improving the model performance by combining water temperature from multi-spectral remote sensing analysis and in-situ measurements. Results of the approach on a eutrophic reservoir in Central Indiana show that with four images of multi-spectral remote sensing data assimilated, the model results oscillate more from the in-situ measurements during the data assimilation period. For validation, the data assimilation has negative impacts on the root mean square error. According to quantitative analysis, more significant water temperature stratification leads to larger deviations. Sampling depth differences for remote sensing technology, in-situ measurements and model output are considered as possible error source.

Meghna Babbar-Sebens, Ph.D., Chair

TABLE OF CONTENTS

1   INTRODUCTION

In recent years, deterioration of water quality in reservoirs that serve as drinking water sources has become one of the major sources of human health risks. Numerical models have been successfully used to simulate the physical, chemical and biological processes within reservoir systems, and predict the risks of contamination (USEPA, 2002; Jin and Ji, 2004; Khangaonkar et al., 2005; Tetra Tech, 2009). Among all the contributing factors that influence the water quality condition in a reservoir system, water column temperature has significant impacts on the distribution, transportation, and interaction of multiple contaminants such as nutrients, micro-algae, etc. This research focuses on modeling hydrodynamics and water column temperature in Eagle Creek Reservoir (ECR), IN, and investigates data assimilation strategies that can incorporate remote sensing and in-situ field measurement data for real time model updates of water temperature.

The research uses the Environmental Fluid Dynamics Code (EFDC) as the main modeling framework for simulating hydrodynamic and water quality in ECR. Although EFDC provides the 3-dimensional simulation of underlying physical, chemical, and biological processes in continuous time scale, there can exist discrepancies between the model output and observations (obtained via remote sensing technology or via in-situ measurements). These errors in the model predictions can arise from inaccurate input, inaccurate model parameters, numerical errors during the computation processes, and real-time variations in the system that are not incorporated in the original calibrated model. A data assimilation procedure can, however, provide means for integrating real-time observed data from variety of monitoring sources to improve a model's prediction

1

accuracy via changes in state variables, model inputs, parameter updates, and/or bias

correction (Moradkhani, 2008). Numerous studies have investigated remotely sensed

datasets to improve predictions of soil moisture (Li et al., 2004; Boussetta et al., 2008;

Zhu et al., 2009), subsurface soil temperature (Olioso et al., 1999), snow cover (Slater

and Clark, 2006) via land surface models, and surface water distributions via

hydrodynamic and water quality numerical models. Study focuses include river

sediments, hydro-meteorological, water quality and sea circulation (Yang and LeDimet,

1998; Seo et al., 2003; Madsen, 2006; Panteleev et al., 2007; Voutilainen et al., 2007;

ADEM/Water Quality Branch, 2008).

Existing methods of data assimilation are based on two types of approaches for

finding the best estimates of state variables, input variables and boundary conditions from

(noisy) observations given a (noisy) model (Walker, 2005). The first approach uses a

"direct observer" and provides a four-dimensional data assimilation scheme, whereas the

second approach uses a "dynamic observer" and provides a sequential data assimilation

scheme. The commonly used direct observer data assimilation approaches include Direct

Insertion (Houser et al., 1998; Robinson and Lermusiaux, 2000), Statistical Correction

(Houser et al., 1998), Nudging (Rizzoli and Young, 1995; Houser et al., 1998), and

Kalman Filter related methods. These "direct observer" methods adjust the model by

continuous update according to the observation in the previous time step. However, most

of these studies assume availability of continuous and sequential measurements, which

might not be easily possible for most water resources problems that use irregularly

available in-situ and/or remote sensing data.

The "dynamic observer" data assimilation approaches adjust the state variables at the beginning of each assimilation window so that model predictions over that time period correspond with the observations. Hence, "dynamic observer" techniques can be posed as optimization problems with strong constraints (variational methods) or weak constraints (dual variational or representer methods). For example, Dente et al. (2008) successfully used a variational method to assimilate ASAR and MERIS satellite data into a wheat model and improved the wheat yield mapping. A cost function is defined to measure the error between model outputs and observations. In order to search for the optimal configuration of the model initial condition and minimizing the cost function, constraints on the variability of initial parameters were set as Gaussian distributed within reasonable ranges. Then the model was reinitialized within this set of initial conditions and the optimum model simulation of wheat yield was obtained. In the research done by (Ines et al., 2006), remote sensing data from two Landsat-5 Enhanced Thermatic Mapper Plus (ETM+) band 6 images were assimilated into a soil-water-atmosphere-plant model using variational data assimilation method. A genetic algorithm (GA) was used in data assimilation to modify model initial conditions and water management optimizations. A variational data assimilation method has also been used to assimilate a sequence of satellite images into a simple transport-diffusion model to simulate the ocean surface current (Korotaev et al., 2008). These approaches are especially useful when observations are available at irregular time intervals, and/or at only specific spatial locations.

Although data assimilation algorithms have been widely applied in environmental studies, few studies related to surface water systems (Panteleev et al., 2007; Korotaev et al., 2008) have used "dynamic observer" data assimilation approaches into the

hydrodynamic models. Among these researches, for example, (Panteleev et al., 2007) assimilated regular observations of flow velocity in the Kara Sea from ADCP instruments. Few previous investigations have been applied with respect of assimilating remote sensing data into numerical models.

The specific goals of this study are:

(1) To develop a 3-Dimensional finite difference numerical model for simulating hydrodynamic processes in ECR in Central Indiana.

(2) To assimilate water surface temperature retrieved from multi-spectral Landsat-5 ETM+ band 6 images into the hydrodynamic model and adjust the model's initial conditions via optimization-based variational data assimilation approach.

(3) To validate the hydrodynamic model by comparing the model results with remote sensing observations, and test the data assimilation efficiency.

2   METHODOLOGY

2.1   Study Area and Data Collection

Figure 1: Geographic location of ECR.

Eagle Creek Reservoir (ECR) is located about 16 km (10 miles) northwest of Indianapolis, Indiana (Figure 1). It was constructed in 1967 by the city of Indianapolis and was initially used for flood mitigation. A water treatment plant was later constructed and put into service in 1976. The treatment plant takes water directly from the reservoir (approximately 10 MGD) and serves primarily as a source of drinking water supply. ECR is a small and shallow reservoir with normal pool surface area 5.1 km^2 and mean depth 5.7 m. It can be separated into three functional areas: the quarry, the northern basin, and the southern basin. The reservoir's northern and southern basins are separated by a land

bridge causeway under 56th street, which allows limited water exchange through an approximately 50 meter opening. The quarry doesn't have a direct connection with the other two basins in ECR and is considered as an isolate feature. Flow in ECR is supplied by tributaries from the upstream Eagle Creek watershed (426 km^2). Four main streams flowing into the reservoir are Eagle Creek, Bush Creek, Fishback Creek and School Branch, with Eagle Creek (mean discharge 4.2m^3/s) being the major contributor of the flow.

Seasonal and short-term temperature changes can lead to the thermal stratification of the reservoir water. Distinct thermoclines that separate the reservoir into warm and cold water zones will prevent the water from mixing. In this case, contaminants tend to accumulate instead of circulating. For example, chlorine tends to accumulate at the bottom of warm water thermoclines (USEPA, 1999). Also, a lack of rainfall with inadequate mixing of fresh and stagnant water, increased algae growth, deterioration of organic matter as the water warms up, and low wind conditions, can call contribute to depletion of DO levels. If a reservoir becomes stratified as a function of temperature, the bottom layer can become deficient in dissolved oxygen (Goodin, 1995). Concentrations of phosphorous, ammonia, iron and manganese, however, are greatly influenced by the presence of oxygen (APEC, 2010).

Bathymetry data for ECR was measured by Center for Earth and Environmental Sciences (CEES), IUPUI. The bottom elevation varied from 223.8m to 240.5m above the sea level. Since there were no existing flow monitoring stations on the major tributaries just north of the reservoir, inflow discharges into the reservoir from the watershed were obtained from a watershed model – Soil and Water Assessment Tool (SWAT) – in this

study. The outflow from the dam was measured at a United States Geological Survey (USGS) gauge Station #03353460 at Clermont (1 km downstream from reservoir). Hourly atmospheric data was obtained from National Climatic Data Center (http://www.ncdc.noaa.gov/). Observations from Eagle Creek Airpark/Airport (53842) station were used for this reservoir region. Hourly solar radiation for 2008 was obtained from Indiana State Climate Office (http://climate.agry.purdue.edu/climate/). The closest station - Throckmorton-Purdue Agricultural Center (TPAC) - reporting solar radiation is located in Lafayette, Tippecanoe County, IN which is 63 miles Northwest from Indianapolis. The evaporation in ECR was not known. However, the local water utility company measured daily evaporation at Carmel, 27 km (17 miles) east of ECR. Data collected by the company indicated an average evaporation value of approximately 5.50 mm/day from June to October, and an average of 4.01 mm/day from November to May (Lobligeois, 2009). Daily pool elevation data for 2008 was obtained from USGS gauge Station #03353450 at ECR located to the east of the dam.

2.2   Simulation Model

Environmental Fluid Dynamics Code (EFDC) (USEPA, 2010) is a public domain, open source, surface water numerical modeling system for simulating hydrodynamics, and water quality in open-surface water bodies. EFDC has been applied to over 100 water bodies in support of environmental assessment and management and regulatory requirements. The EFDC model solves the three-dimensional, vertically hydrostatic, free surface, turbulent averaged equations of motions for a variable density fluid. The model uses a stretched or sigma vertical coordinate and Cartesian or curvilinear, orthogonal

horizontal coordinates. The hydrodynamic model also solves dynamically coupled transport equations for turbulent kinetic energy, turbulent length scale, salinity, and temperature.

The 3-D stretched sigma grids implemented with the EFDC use the following transforming function to calculate an adjusted vertical coordinate, Z, from the bottom elevation and water surface elevation:

$$Z = (Z^* + h) \Big/ (\zeta + h) \tag{1}$$

In Equation 1, Z^* = original physical vertical coordinate, h = bottom elevation and ζ = water surface elevation. After the physical vertical coordinate system is stretched, the total depth is evenly distributed into equal depths of individual layers, for all the X-Y grid locations within the research domain.

The continuity equation used in EFDC is given by Equation 2, in which H = water depth, u and v = horizontal velocity components in x and y direction respectively, w = vertical velocity component in z direction; Q_H = the volumetric source and sink term concerning rainfall, evaporation and infiltration. The conservation of momentum equations are given in Equations 3 and 4, in which f = Coriolis factor, p = the water column hydrostatic pressure; p_{atm} = the kinematic atmospheric pressure; Av = vertical turbulent momentum diffusion coefficients, and Qu and Qv = momentum source-sink terms.

$$\frac{\partial H}{\partial t} + \frac{\partial (Hu)}{\partial x} + \frac{\partial (Hv)}{\partial y} + \frac{\partial w}{\partial z} = Q_H \tag{2}$$

$$\partial_t(Hu) + \partial_x(Huu) + \partial_y(Hvu) + \partial_z(wu) - fHv = -H\partial_x(g\zeta + p + p_{atm}) +$$
$$(\partial_x h - z\partial_x H)\partial_z p + \partial_z(H^{-1}A_v\partial_z u) + Q_u \tag{3}$$

8

$$\partial_t(Hv) + \partial_x(Huv) + \partial_y(Hvw) + \partial_z(wv) - fHu = -H\partial_y(g\zeta + p + p_{atm}) +$$
$$(\partial_y h - z\partial_y H)\partial_z p + \partial_z(H^{-1}A_v\partial_z u) + Q_v \tag{4}$$

In the transport equations for salinity and temperature (Equations 5 and 6) the source and sink terms are given by Q_S and Q_T, which consist of subgrid scale horizontal diffusion and thermal sources and sinks, and A_w is the vertical turbulent diffusivity.

$$\partial_t(HS) + \partial_x(HuS) + \partial_y(HvS) + \partial_z(wS) = \partial_z(H^{-1}A_v\partial_z S) + Q_S \tag{5}$$

$$\partial_t(HT) + \partial_x(HuT) + \partial_y(HvT) + \partial_z(wT) = \partial_z(H^{-1}A_v\partial_z T) + Q_T \tag{6}$$

The water surface and bed boundary conditions for heat transport are given by Equations 7 and 8:

For water surface:

$$-\frac{A_w}{H}\partial_z(T) = \frac{(J_b + J_c + J_e)}{\rho C_{pw}} \tag{7}$$

For bed:

$$\partial_t(H_b T_b) = \frac{I_b}{\rho_b} - C_{hb}\frac{\rho C_{pw}}{\rho_b C_{pb}} {}^2\sqrt{(u_{bl}^2 + v_{bl}^2)}(T_b - T_{bl}) \tag{8}$$

Short wave solar radiation at the bed is defined as:

$$\frac{I_b}{I_s} = re^{-\beta_f H} + (1 - r)e^{-\beta_s H} \tag{9}$$

where J_b = net long-wave back radiation; J_c = convective heat transfer; J_e = evaporation heat transfer; C_{pw} = specific heat of water; H_b = active thermal thickness of the bed; T_b = bed temperature; I_b = short-wave solar radiation at the bed; ρ_b = bed density; C_{pb} = specific heat of the water-solid bed mixture; C_{hb} = dimensionless convective heat exchange coefficient; T_{bl} = bottom layer water temperature; I_s = solar radiation at the

water surface, r = distribution factor; β_f and β_s = fast and slow-scale attenuation coefficients (Caliskan, 2008).

Eight equations (2-9) provide a closed system for the variables u, v, w, p, ζ, ρ, S, and T. The vertical turbulent viscosity and diffusivity and the source and sink terms are also specified (Hamerick, 1992).

2.3   EFDC Grid Generation

Multiple grid sizes and time steps (Appendix C) for representing the physical system were explored in the study assure the accuracy of model results as well as the efficiency of the model. The preliminary condition for 3- time- level numerical scheme is that the model grid size and time step should meet the Courant-Friedrichs-levy condition (CFL condition) which is a necessary condition for convergence while solving certain partial differential equations (usually hyperbolic PDEs). CFL condition is expressed as Equation 10:

$$\gamma = U \frac{\Delta t}{\Delta x} \tag{10}$$

where U = velocity, Δt = time step, Δx = cell size. The necessary restriction for grid size and time step to ensure numerical convergence and stability is $\gamma < 1$.

The final grid setup for the numerical model consisted of expanding grids with minimum grid size 40m to maximum 60m are used. The expanding factor of 1.005 was chosen to expand grid sizes from the focal point, which was the water intake. This location was chosen to accurately simulate the most complex flow condition happening close to the causeway under the 56th Street land bridge, and the drinking water intake. A total of 2401 grid cells were developed to represent the physical domain in the modeling

domain. This grid best represented the shape of the reservoir shoreline compared to other grids. Since EFDC only recognizes flow through cell faces, cells connecting with each other by corners do not exchange any mass or momentum across the corners (Figure 2). For this grid system, the time step of the finite difference model was set up to two seconds considering the model stability and the computational burdens.
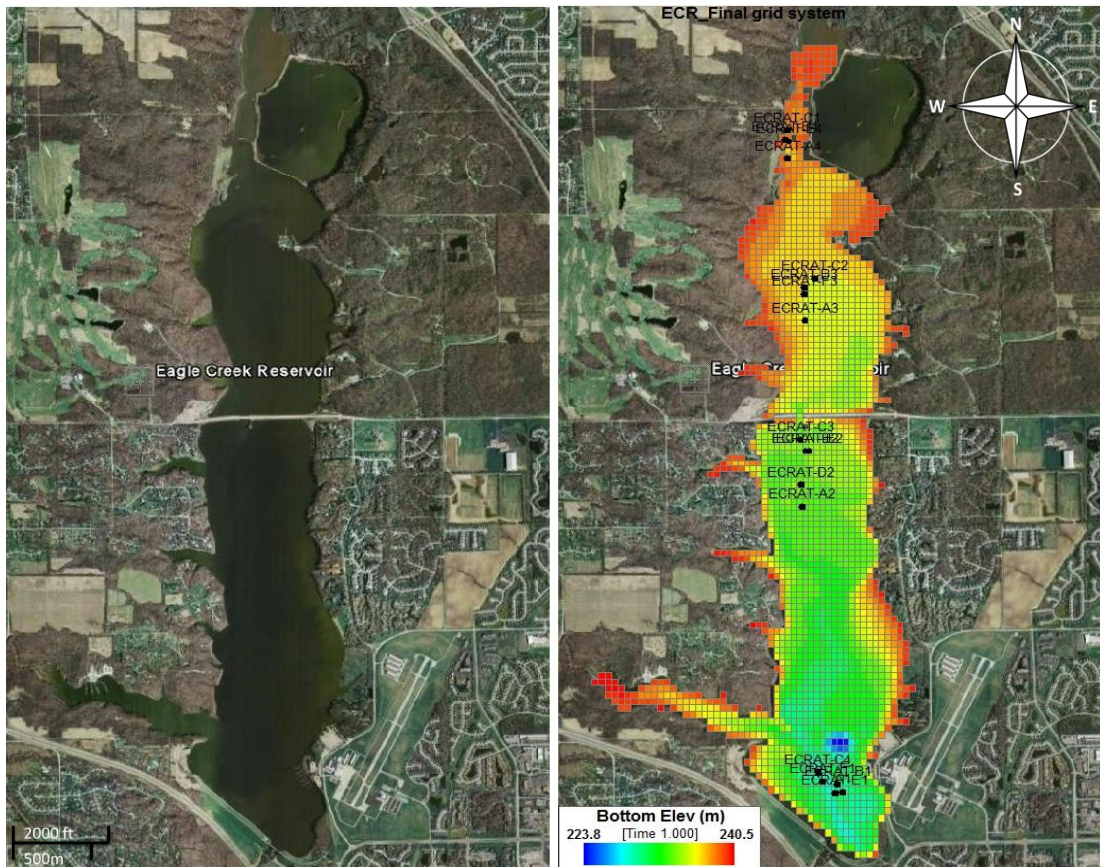
Figure 2: ECR morphological conditions (left) and final grid system with 2008 sampling locations (right).

2.4   Initial/Boundary Conditions

The depth of water through the reservoir was measured by CEES (Center for Earth and Environmental Science, IUPUI) based on in-situ measurements with sonar equipments. Bottom elevation was based on the bathymetry data, and the initial water surface elevation (or pool elevation (PE)) for the model was chosen to be 240.56 m according to USGS (station near the dam) measurement on January 1st, 2008. A uniform initial water temperature of 3.4˚C throughout the reservoir was assumed for the model, based on the measurements at Mill Creek USGS gauge station near Manhattan, IN on January 1, 2008.

Hydrodynamic boundary conditions of ECR included (a) time-series inflow discharge from twelve tributaries of Eagle Creek simulated by the Soil and Water Assessment Tool - based watershed model (Figure 3), (b) outflow discharge through the water intake and ECR dam, (c) wind speed and direction, and (d) atmospheric data including precipitation and evaporation. All data were appropriately formatted and imported into the EFDC modeling system. Figure 3 (right) shows the boundary where the watershed tributaries join the EFDC modeling grid for ECR.
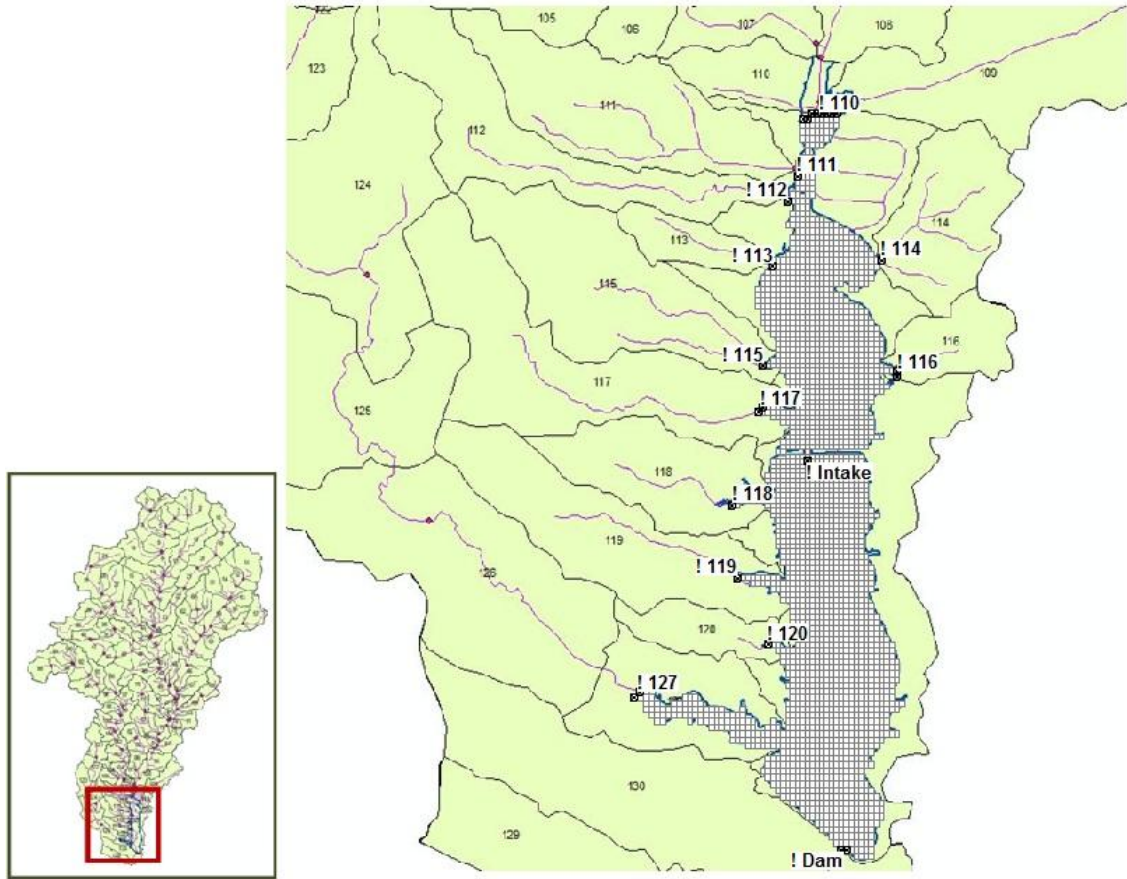
Figure 3: Eagle Creek Watershed tributaries that interface with the numerical modeling grid of ECR.

2.5 Data Assimilation Algorithms

The overview of the data assimilation process is shown schematically in Figure 4. The main objective of the data assimilation process was to change the values of initial conditions of temperature in the reservoir at the beginning of the simulation, so that there was minimum error between model predictions of temperature and observations at a future time. The EFDC model used the set of assumed initial conditions, monitored and modeled (e.g., those from the SWAT hydrologic model) boundary conditions, and calibrated model parameters to predict the 3-D spatial and temporal distribution of

modeled physical variables. A cost function that estimates the error between predictions

and observations at specific time periods when observations were obtained was then used

to direct an optimization algorithm to modify the assumed initial conditions on the

reservoir, until an optimum set of initial conditions are obtained.
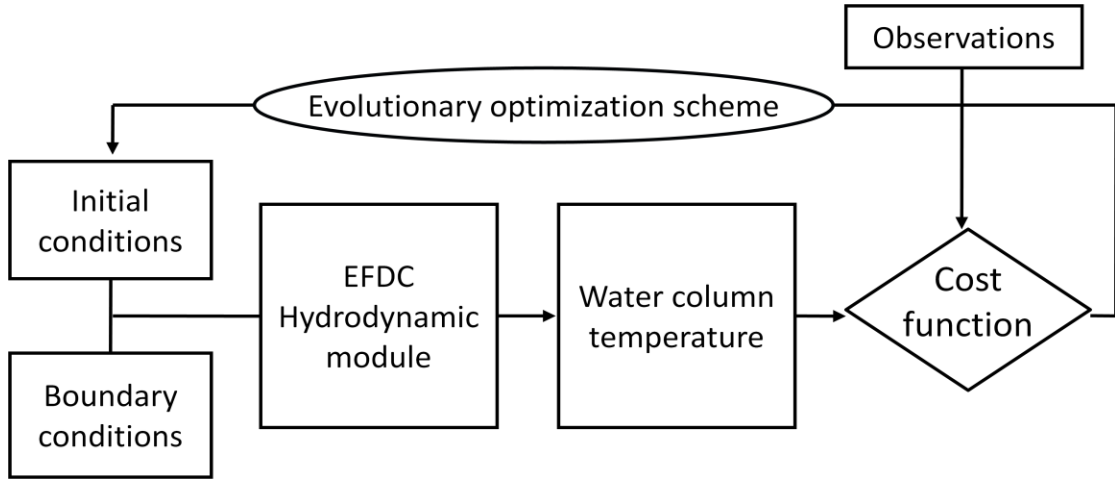
Figure 4: General scheme of remote sensing data assimilation.

Since the main objective of this study was to examine the advantages of using

remote sensing observations for updating hydrodynamic and water quality numerical

models, multispectral data were obtained from the Landsat 5 satellite. The spatial

resolution for ETM+ band 6 images was re-sampled from 120m * 120m to 30m * 30m.

Four of the ETM+ band 6 images obtained from the satellite were used to convert

spectral radiance to water surface temperature based on the Planck's law (Equation 11).

In Planck's law, the longwave radiation emitted from the land surface was in proportion

to its temperature as:

$$T(\lambda) = \frac{K_2}{\ln\left(\frac{K_1}{L(\lambda)} + 1\right)} \tag{11}$$

where λ is the wave length; K_1 and K_2 are the calibration constants as 607.76 Watts/(m^2 *

ster * μm) and 1260.56 Kelvin respectively (Goodin, 1995; NASA, 2009); L is the

spectral radiance in watts/(m^2 * ster * μm). These pre-launch calibration constants from

the empirical models are used under the assumption that the downwelling radiance and

atmospheric transmissivity are constant in space throughout the study area, and thus are

applied to calculate surface temperature for each image pixel (Quattrochi, 2004). Studies

using this method found the RMSE in water surface temperature to be less than 1K of

retrieving land surface temperature (Sobrino et al., 2004). This error is even smaller for

surface water systems because of their better homogeneity in temperature.

To assimilate data obtained from the Landsat ETM+ images, 300 random

locations were identified in the reservoir (Figure 5), where the model results and optical

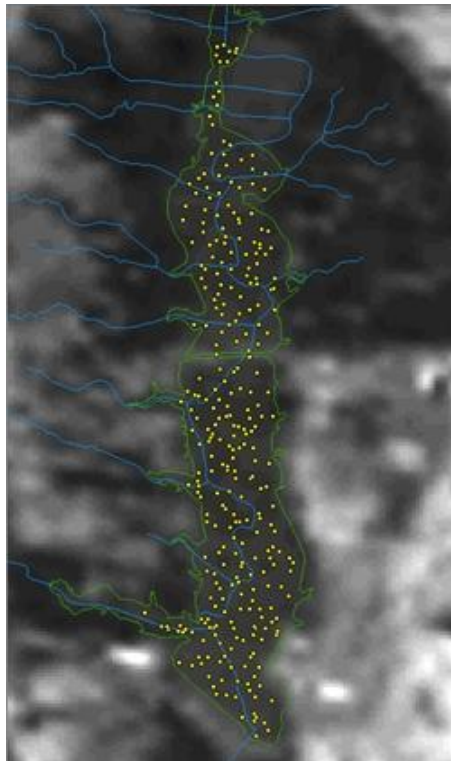observations were used to evaluate the model performance.

Figure 5: Random locations identified in the reservoir with ETM+ image on Aug 7, 2008.

The finite difference model computed and generated results in 2401 cells distributed over five vertical layers within ECR. Since the remotely sensed observations are most applicable to the water surface conditions, the very top layer of the finite difference grids were chosen for comparing model predictions to the corresponding remote sensing results.

A variational observer data assimilation algorithm (Figure 6) was used in this study to incorporate the remote sensing observations within the simulation model. The error term from observations was omitted under the assumption (1) Gaussian distribution of observation errors; (2) The EFDC model and the remote sensing data retrieving algorithm were not biased (Dente et al., 2008). The data assimilation window was defined as the time period from the initial condition to the last observation time (Figure 6). The cost function (Figure 4) was estimated using the relative root mean square error (RRMSE) calculated from both remote sensing outputs and model outputs in the topmost layer of all 300 random locations (Equation 12). RRMSE calculated for each of the four observation days was then equally weighted and summed to obtain an overall RRMSE for the cost function (Equation 13).

$$RRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(X_o - X_m)^2}{n}}}{\overline{X_o}} \tag{12}$$

$$Cost\ funtion = \frac{\sum_{1}^{j} RRMSE_j}{j} \tag{13}$$

Where X_o is values of the field observed model parameter; X_m is the model outputs; n is the number of observations; $\overline{X_o}$ is the mean value of the field observations; j is the number of remote sensing observations.
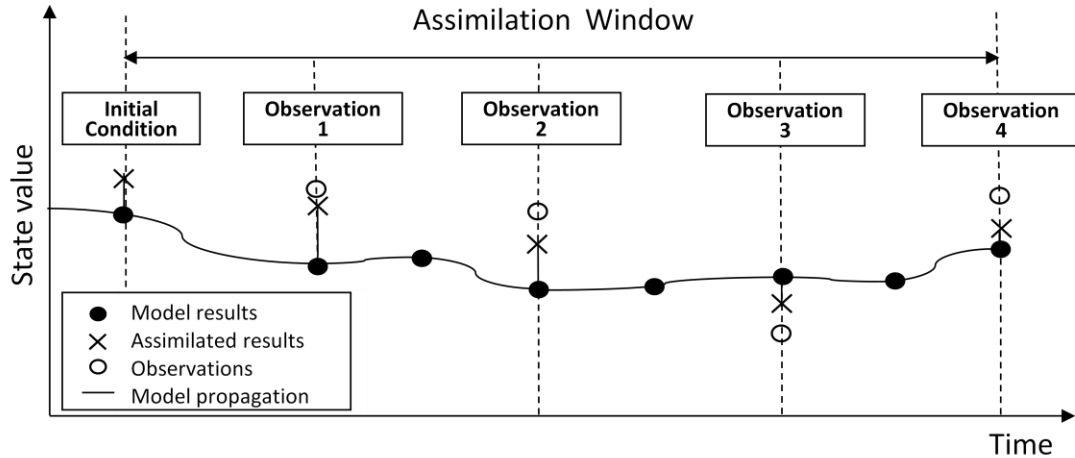
Figure 6: Scheme of time series evolutionary data assimilation process.

A single objective genetic algorithm (Goldberg, 1989) was used in this study for minimizing the cost function (Equation 12). The decision variables were defined as the percentage change in the initial conditions for each vertical layer. The first remote sensing image came at 219th day (August 7th) of 2008. The data assimilation window was started from 213th (August 1st) day to allow a warm-up period for EFDC. In the first GA loop, the initial range of percentage change in water temperature for all the five vertical layers was defined within -20 to 20. In the following loops, this range was further changed for each vertical layer according to the best cost function results in the previous loop. In total five loops were implemented to the GA algorithm with nine possible values of decision variable evenly taken within the decision variable ranges (Table 1). The population size was set to be 16; the crossover and mutation rate were 0.90 and 0.05 respectively. The GA code was run for 10 generations.

Table 1: GA decision variable setup ranges for optimization.

	Layer1	Layer2	Layer3	Layer4	Layer5
Loop1	(-20,20)	(-20,20)	(-20,20)	(-20,20)	(-20,20)
Loop2	(-15,5)	(-20,0)	(-20,0)	(-20,0)	(-15,5)
Loop3	(-20,0)	(-25,0)	(-25,0)	(-25,0)	(-20,0)
Loop4	(-30,-15)	(-30,-15)	(-35,-20)	(-35,-20)	(-30,-15)
Loop5	(-55,-25)	(-60,-30)	(-60,-30)	(-60,-30)	(-55,-25)

3   MODEL APPLICATION AND RESULTS

3.1   Model Calibration

3.1.1   Water surface elevation and mass balance

The period January 1st, 2008 to July 31st, 2008 were used for the calibration
process, just before the first set of remote sensing observations were available. Figure 7
shows the discrepancy between measured pool elevation in the reservoir for the year
2008 and the pool elevation calculated via mass balance of SWAT-modeled tributaries
inflows, monitored water intake and dam outflow. The root mean square error (RMSE)
for the calculated pool elevation in Figure 7 was estimated to be 1.036 meters.
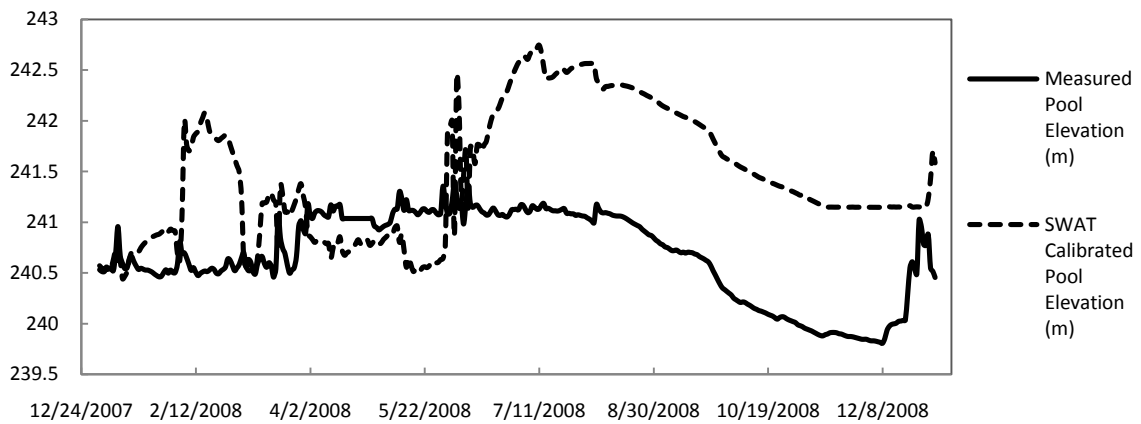
Figure 7: Field measurements vs. SWAT calibrated water surface elevation in 2008.

These discrepancies arose due to errors in the estimated tributaries inflows, which
indicated need for further adjustment of inflows in order to better match the observed
pool elevation. Also, since such large discrepancies in water surface elevation could lead
to significant mismatches in water column temperature vertical profiles, a further
adjustment in SWAT model outputs of flow discharges was made in two phases.

Phase I – Adjustment of SWAT-modeled discharge in tributaries

The first step in adjusting the SWAT inflows used for the EFDC model included calculation of daily net storage in the reservoir from the inflow and outflows based on the following mass balance equation 14:

Daily net storage in reservoir at the end of the day i = Total daily inflows from the tributaries on day i – daily outflow in the water intake on day i – daily outflow in the dam on day i   (14)

Once the daily net storage was estimated, it was added to the volume of the water in the reservoir to estimate the volume of the reservoir at the beginning of the next day (Equation 15).

Volume in reservoir on day i+1 = Volume in reservoir on day i + Daily net storage in reservoir at the end of the day i   (15)

A rating curve (Appendix E) was then developed using the bathymetry data, to estimate relationships between reservoir volume and pool elevation. This rating curve provided means to estimate the reservoir volume at any day based on the pool elevation measured on that day. The difference between the calculated and measured water volume in the reservoir was weighted according to the tributary discharge for that day, and was subtracted from the daily tributary inflows (Equation 16).

Adjusted discharge for tributary k on day i= Original discharge for tributary k – |calculated water volume – measured water volume| on day i * discharge weight for tributary k   (16)

After tributary inflows were adjusted, the new net storage of water within the

reservoir and the corrected total volume on day i were calculated. At the beginning of day

i+1, the initial volume was the summation of corrected water volume in the day i and the

net storage from inflow and outflow. The total difference between measured and

calculated water volume was again going through the same correction step, generating

the corresponding tributary flows for day i+1. This adjustment was applied to all daily

inflows until the end of the model simulation period. The schematic procedure of

computation is shown in Figure 8. Though this adjustment produced the water surface

pool elevation identical to the measured data (Figure 9), it also resulted in some negative

values in the adjusted tributary flows that were corrected in the next step.
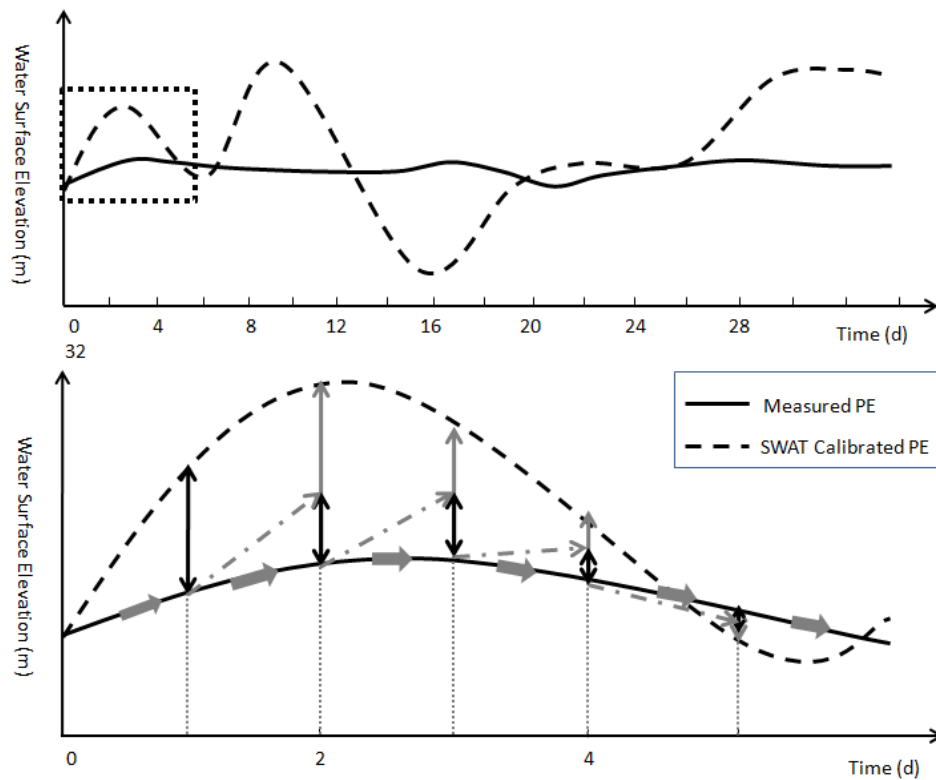
Figure 8: Water surface elevation as a result of flow adjustment. (The bottom figure is an

enlarged sub-section of the top figure)
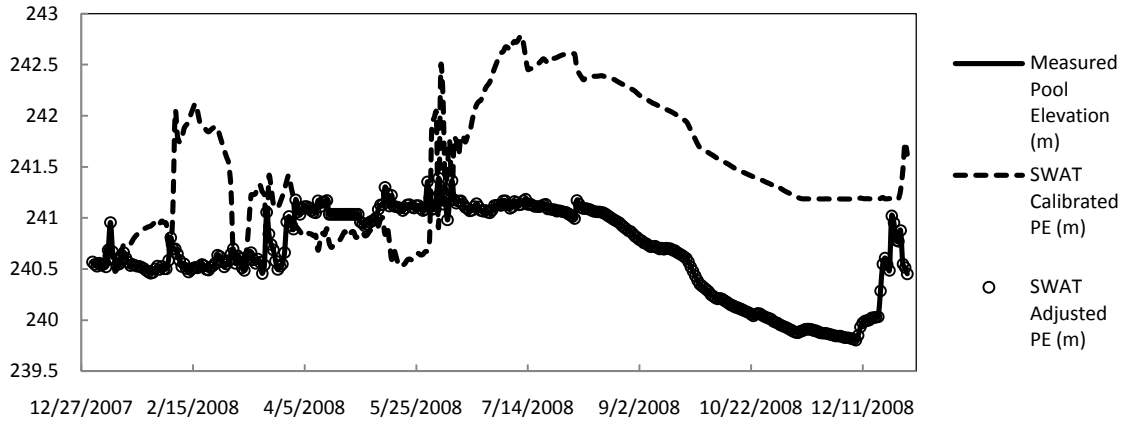
Figure 9: Results of adjustment in water surface elevation.

Phase II – Correction of flows

In order to avoid the negative tributary flows, all the negative flows were replaced by a very small positive discharge of $0.00001 m^3/s$. This resulted in a small discrepancy with 0.07m RMSE between the adjusted and measured water surface elevation (Figure 10). This error after flow correction was within the reasonable range in scale.
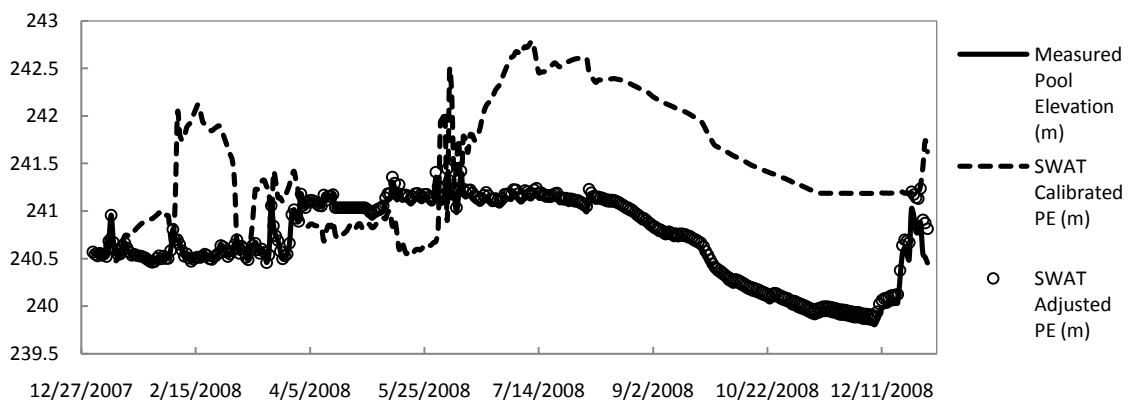
Figure 10: Results of water surface elevation after negative flow correction.

Phase III – Pool elevation simulation using calibrated SWAT flow

The adjusted and corrected SWAT tributary inflows were then used as inputs into the reservoir model (EFDC) and the model was run for the period from January 1st, 2008 till July 31st, 2008. Figure 11 compares the EFDC simulated water surface elevation with the measured pool elevation. Since, EFDC simulates at much small time step than the daily time-step SWAT model, the EFDC output at the end of each day was taken as the corresponding simulated value for the water surface elevation on that day in Figure 11. Comparison of the modeled results and the field measurements during the first seven months in 2008 produced a root mean square error (RMSE) of 0.0287 meter, which was 0.087 meter over the 0.2 meter instrumental accuracy in measuring the bathymetry (Lobligeois, 2009).
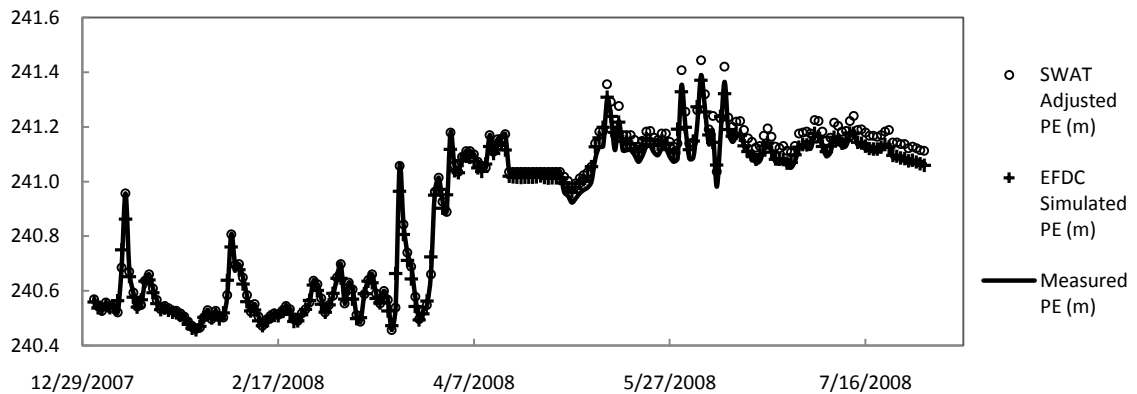
Figure 11: Calibrated water surface elevation (purple) based on adjusted flows from SWAT.

3.1.2   Water column temperature

Other than water surface elevation, water column temperature was calibrated in full hydrodynamics. The in-situ measurements were collected at 54 different (X, Y) locations throughout the reservoir with multiple measurements in several depths at each location in 2008. The YSI probes which were used to measure the water column temperature at these sampling locations reported an instrumental error of ±0.15 ^{o}C (YSI, 2011). These measurements were used for model calibration and validation. During the model calibration period, a series of calibration parameters were adjusted using EFDC to achieve the best agreement with the in-situ measurements of water column temperature (Table 2). In this research, water temperature calibration took 23 sampling events between May 22, 2008 and July 30, 2008.

Table 2: Water temperature calibration parameters.

Water temperature	Calibrated Value
Clear water light extinction coefficient (1/m)	7
Solar radiation input/internally computed	False
Heat transfer coefficient between bed and water column	0.0000005
Evaporation transfer coefficient	0
Min fraction of solar radiation absorbed in the top layer	1
Initial bed temperature (^{o}C)	10
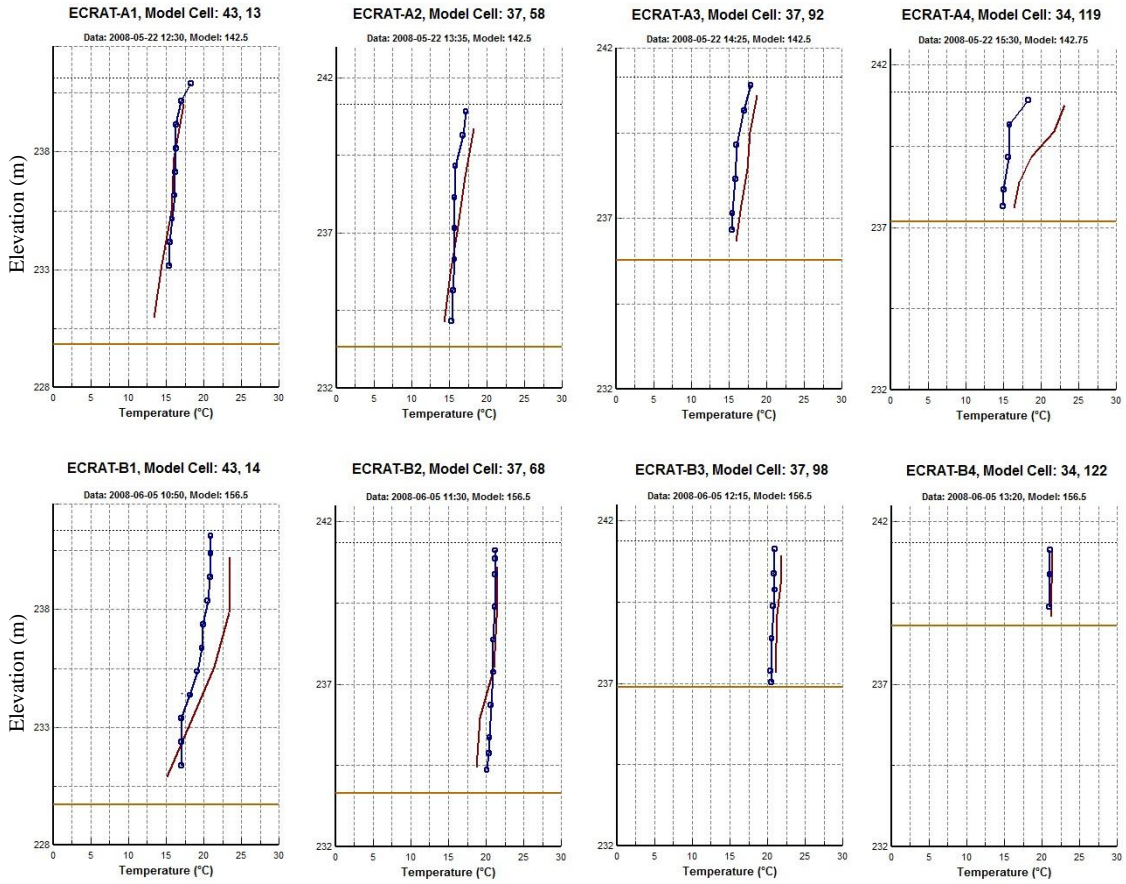
Figure 12 (a): Calibrated temperature vertical profiles for ECR.
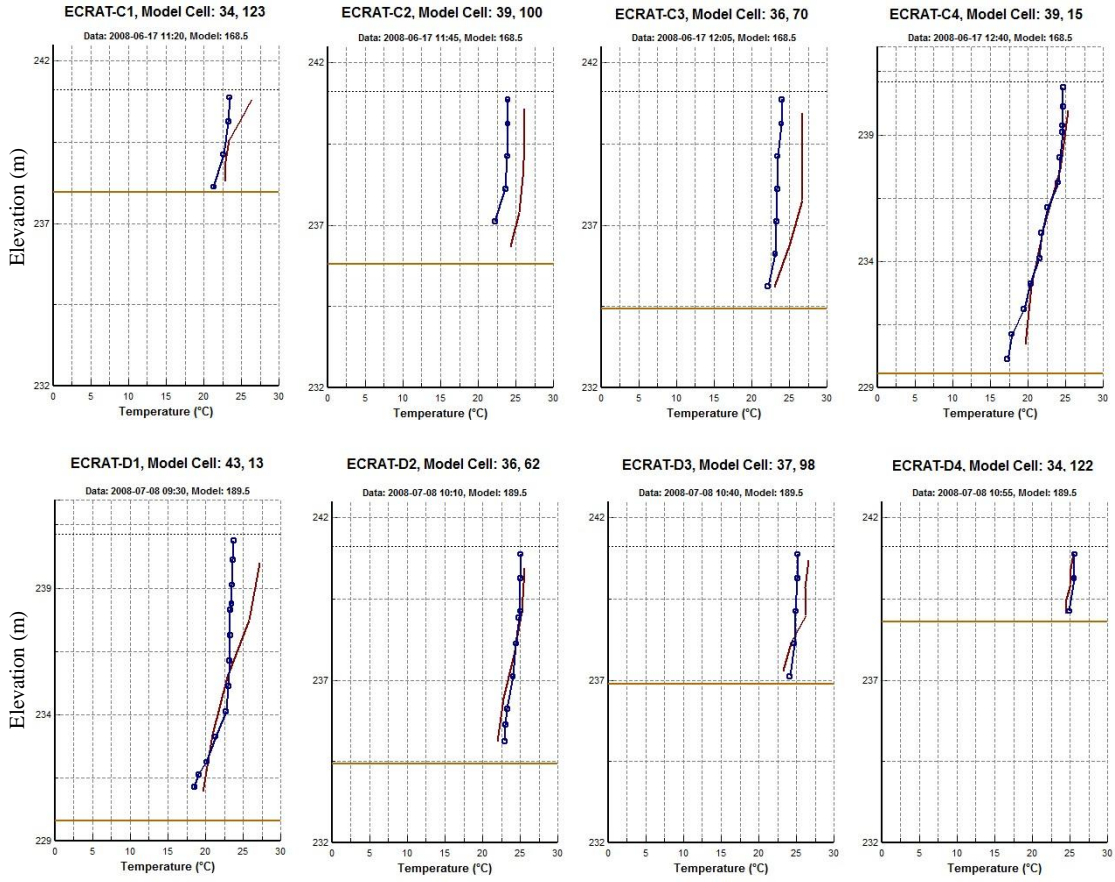
Figure 12 (b): Calibrated temperature vertical profiles for ECR.
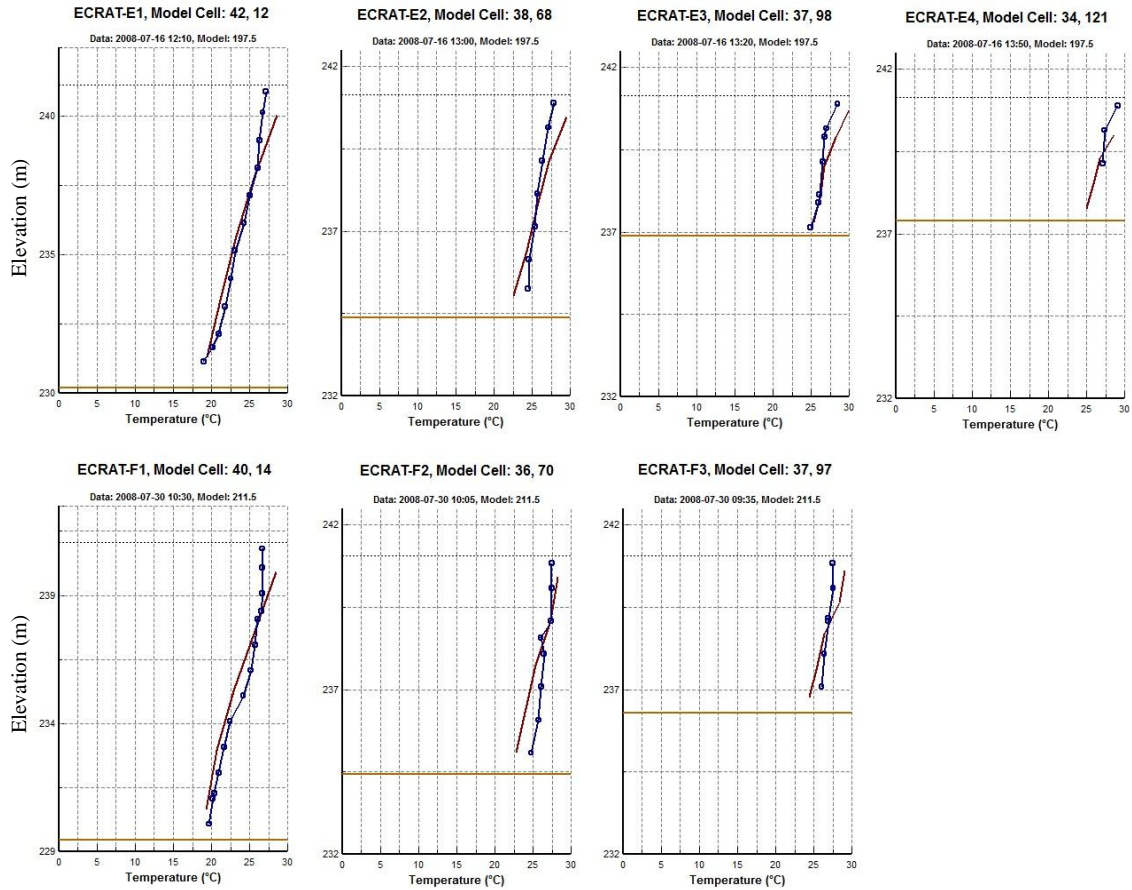
Figure 12 (c): Calibrated temperature vertical profiles for ECR.

The temperature calibration results are shown in Figure 12 (a-c). The densely dashed lateral lines show the water surface elevation of each corresponding water column; the solid lateral lines represent the bottom elevation. It indicates very good correlation between the measured (dotted curves) and the modeled (smoothed curves) for water column temperature. Statistical analysis shows an overall RMSE of 1.279 ℃, which is within the satisfaction range for water temperature simulation using EFDC (Table 3). From the statistical view, the water temperature simulation has the most significant percentage error at monitoring Station ECRAT_A4. This phenomenon can be explained because this station is located at the upstream riverine area of ECR, and at the downstream of the confluence of an unnamed tributary with the reservoir. In summer, it

is assumed that water around this shallow region of the reservoir has higher temperature than nearby tributary flows from the watershed. The tributary flows bring cooler water from the watershed into this area, where more stable and warmer water conditions have been created. However, due to the lack of data from the watershed model, the reservoir boundary condition for water temperature was determined based on the USGS gauge station #03354000 at White River near Centerton, IN (approximately 25 miles south from ECR). Therefore, the modeled temperature has high deviation from measured temperature at this station.

Table 3: Statistical review of water column temperature calibration results.

Station ID	Date/Time	# Pairs	RMSE (^{o}C)
ECRAT-A1	22-May-08	9	0.574
ECRAT-A2	22-May-08	8	0.955
ECRAT-A3	22-May-08	6	1.175
ECRAT-A4	22-May-08	5	3.931
ECRAT-B1	5-Jun-08	11	2.176
ECRAT-B2	5-Jun-08	10	0.881
ECRAT-B3	5-Jun-08	7	0.701
ECRAT-B4	5-Jun-08	5	0.303
ECRAT-C1	17-Jun-08	5	1.807
ECRAT-C2	17-Jun-08	5	2.368
ECRAT-C3	17-Jun-08	8	2.441
ECRAT-C4	17-Jun-08	13	0.918
ECRAT-D1	8-Jul-08	13	2.099
ECRAT-D2	8-Jul-08	9	0.564
ECRAT-D3	8-Jul-08	5	1.101
ECRAT-D4	8-Jul-08	4	0.401
ECRAT-E1	16-Jul-08	12	0.892
ECRAT-E2	16-Jul-08	7	1.249
ECRAT-E3	16-Jul-08	7	1.114
ECRAT-E4	16-Jul-08	3	1.143
ECRAT-F1	30-Jul-08	14	0.990
ECRAT-F2	30-Jul-08	9	1.156
ECRAT-F3	30-Jul-08	6	1.024
Composite Statistics		185 (sum)	1.279 (average)

With the calibrated water column temperature provided by EFDC calibrated water column temperature, a remote sensing data assimilation approach was implemented during the simulation time period from August 1st to October 10th, 2008.

3.2   Remote Sensing Data Assimilation

Before data assimilation process was applied, three assumptions are made: (1) during the data assimilation procedure, only two sets of data were known - the model outputs and the remotely sensed images; (2) the in-situ measurements was considered as the natural "truth"; (3) the remotely sensed data sets performed in better accuracy than the model outputs, compared with the "truth" (i.e. in-situ measurements). The error was minimized between the remote sensing and model results, and the data assimilation results were evaluated in comparison with in-situ measurements. In total four remote sensing images during 2008 (August 7th, August 23rd, September 24th, October 10th) were applied in data assimilation for water temperature concentration.

During the remote sensing data assimilation processes, EFDC was set up in a restart mode with the restarting time at August 1st, which was the 213th day of the simulation time period. The EFDC model was reinitialized with the water temperature initial condition adjusted at 213rd day of 2008. During this process of reducing the variational range of the decision variables, it was discovered that the decreasing water temperature values in initial condition provide better results, which were closer to the remote sensing observations. Also, the improvements in model performance with adjusting the model initial conditions reduced rapidly with model time-series propagation within the assimilation window.

The results in assimilating water temperature data from the TM images successfully reduced the error fitness function from 20.979% to 18.568%. Figure 13 shows comparison between the model computed surface layer water temperature at 300 random locations (circles) and the value retrieved from Landsat ETM+ images, on the four observation days in 2008. The temperature outputs after data assimilation application showed better consistency with remote sensed observations. These results indicated the well performance of the evolutionary optimization scheme in the data assimilation process.
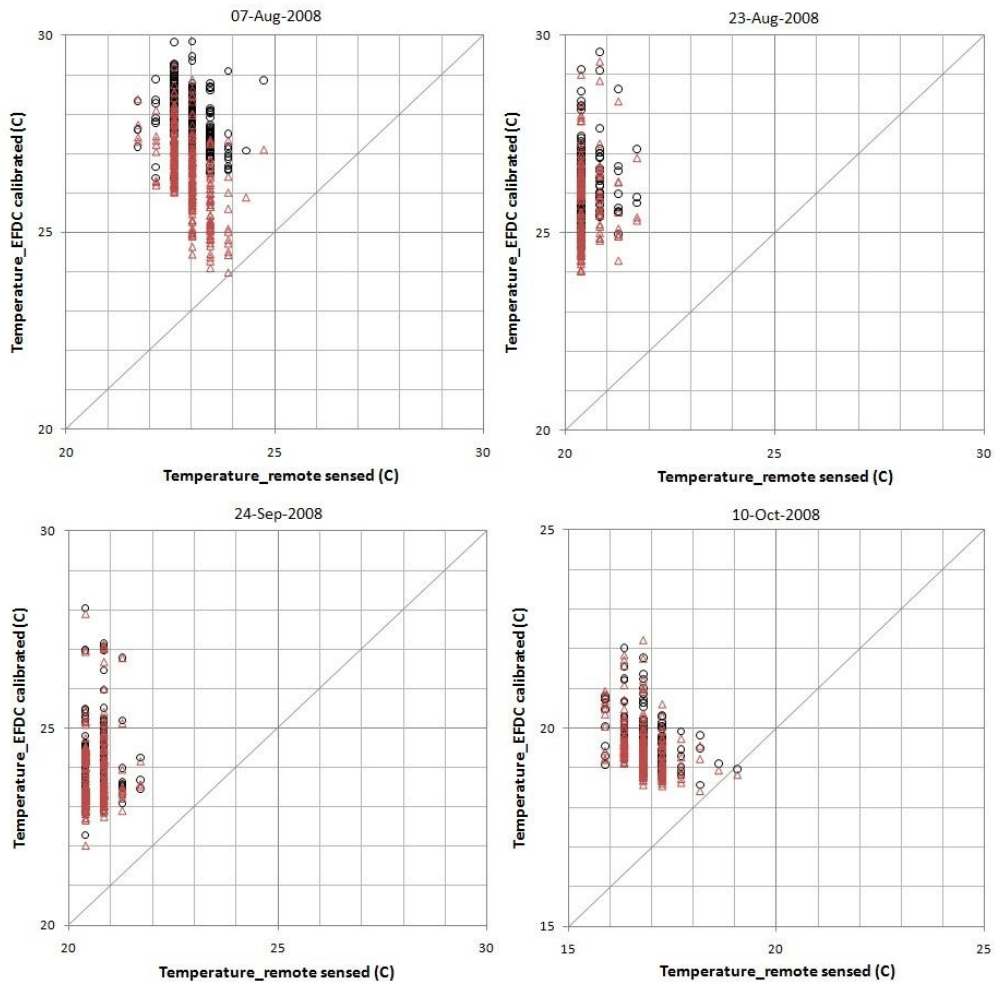
Figure 13: Water temperature data assimilation results: Circles - EFDC generated temp at 300 points without DA adjust; triangles - DA adjusted tempt at 300 points.

The errors were tested between modeled and in-situ measurements before and after data assimilation. The run with error fitness function as 18.568% from GA computation had decision variables as (-25, -25, -25, -12.5, -2.5) for layer one to five (i.e. bottom to surface layer). The comparisons are shown in Table 4.

Table 4: Statistical review of remote sensing data assimilation resulting on water column temperature.

Station	Date/Time	# Pairs	RMSE (^{o}C)	
			EFDC calibrated	DA generated
ECRAT-G1	14-Aug-08	11	0.925	3.420
ECRAT-G2	14-Aug-08	8	1.158	3.081
ECRAT-G3	14-Aug-08	6	1.150	0.897
ECRAT-H1	20-Aug-08	12	2.440	4.628
ECRAT-H2	20-Aug-08	7	2.325	3.913
ECRAT-H3	20-Aug-08	6	2.528	3.383
ECRAT-H4	20-Aug-08	4	2.056	2.605
ECRAT-I1	27-Aug-08	14	2.043	3.872
ECRAT-I2	27-Aug-08	8	1.861	2.900
ECRAT-I3	27-Aug-08	7	1.581	1.576
ECRAT-I4	27-Aug-08	5	4.380	3.994
ECRAT-J1	3-Sep-08	7	2.247	2.261
ECRAT-J2	3-Sep-08	8	1.570	1.862
ECRAT-J3	27-Aug-08	7	2.201	1.653
ECRAT-J4	3-Sep-08	16	2.595	4.024
ECRAT-K1	16-Sep-08	13	1.641	2.677
ECRAT-K2	16-Sep-08	7	0.493	1.043
ECRAT-K3	16-Sep-08	6	0.157	0.490
ECRAT-K4	16-Sep-08	4	1.108	0.936
ECRAT-L1	30-Sep-08	11	1.788	2.505
ECRAT-L2	30-Sep-08	8	0.659	1.089
ECRAT-L3	30-Sep-08	6	1.131	1.028
ECRAT-L4	30-Sep-08	4	2.672	2.617
Composite statistics		185(sum)	1.781(average)	2.692(average)

As the genetic algorithm propagated after a number of generations, the value of error fitness function keeps reducing. Four 'individuals' with their error fitness function results were picked up to indicate the error propagation between the genetic algorithm optimized water column temperatures and the in-situ measurements (Table 5). An increase in the error between the data assimilated model results and in-situ measurements was observed while the error between remotely sensed and model results was reducing. This trend indicates that the remote sensing retrieved water temperature is forcing the model farther away from the in-situ measurements, which is assumed to be the "truth".

Table 5: Optimization results versus corresponding model accuracy in simulating in-situ conditions.

	RS vs.model	model vs. in-situ measurements	
	RRMSE	RelError (%)	RMSE (^{o}C)
Original:	**0.210**	**6.637**	**1.781**
After DA:	0.196	9.017	2.421
	0.193	9.143	2.451
	0.188	9.765	2.614
	0.187	10.065	2.692

3.3   Model Validation and Statistical Analysis

The final phase involved model validation using calibrated model and in-situ water column data for the period of October 10th, 2008 to December 31st, 2008. The main objective of validation phase was to test the model performance for water surface elevation and water temperature predictions with model parameters adjusted through calibration. A statistical comparison between model prediction and measured parameters

in 2008 was implemented after validation. This comparison would serve for quantitative analysis of model accuracy in the research domain of ECR, Indiana.

The water surface elevation validation results is shown in Figure 14. Statistical results indicate that EFDC successfully predict the water surface elevation with an RMSE of 0.085 meters between the modeled and USGS measurements. The validation showed good correlation between these two set of data with an R^2 of 0.97.
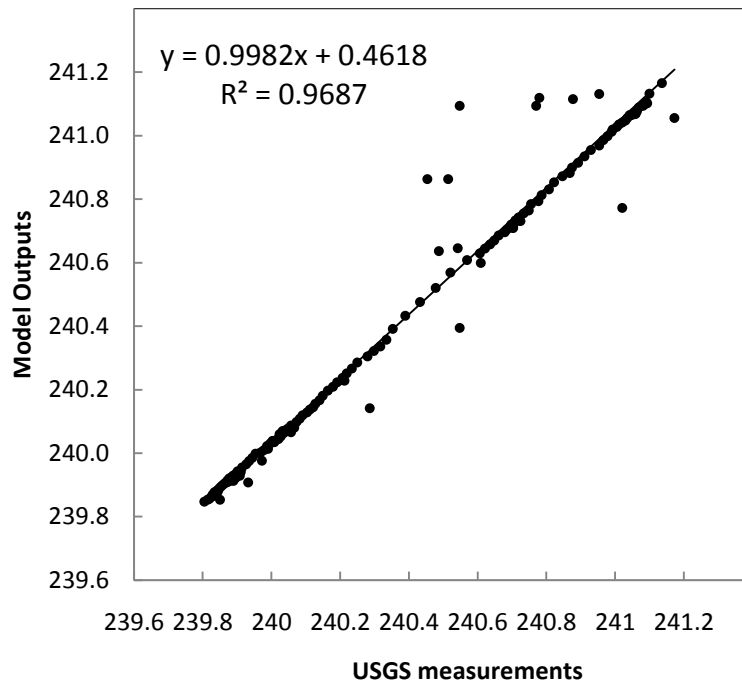
Figure 14: EFDC water surface elevation validation results.

Meanwhile, the model was validated based on the adjustments that are implemented during the data assimilation period. This validation results are compared with those without data assimilation algorithm is applied. The effects of water column temperature data assimilation on the model prediction capability are tested for this study case. It is shown in Table 6 that the RMSE between model outputs and the in-situ measurements increased from 0.960 $^{\circ}$C to 0.962 $^{\circ}$C after data assimilation is applied.

Table 6: Statistical review of water column temperature validation.

| Station ID | Date/Time | # Pairs | RMSE (^{o}C) | |
			EFDC outputs	DA generated
ECRAT-M1	16-Oct-08	10	1.176	0.891
ECRAT-M2	16-Oct-08	6	0.441	0.532
ECRAT-M3	16-Oct-08	4	0.827	0.731
ECRAT-M4	16-Oct-08	3	0.408	0.342
ECRAT-N1	28-Oct-08	10	1.354	1.466
ECRAT-N2	28-Oct-08	6	0.755	0.857
ECRAT-N3	28-Oct-08	4	0.881	0.936
ECRAT-N4	28-Oct-08	2	1.701	1.801
Composite Statistics		45 (sum)	0.960 (average)	0.962 (average)

3.4   Comparison between Remote Sensing Data and In-situ Measurements

The preliminary data assimilation results proved that there exist significant distances between the in-situ measurements and the remote sensing data, whose values were generally higher and lower than the model predictions respectively. These differences might come from random or unknown human error, which could not be easily estimated. Considering the difficulty of comparing the two sets of data which were taken at different time snapshots, four physical regions were identified within the reservoir. A number of in-situ sampling stations were concentrated within each region, and different regions were significantly apart from each other. Region A and B were located in the northern basin of the reservoir, and Region C and D were located in the southern basin (Figure 15). Remote sensing data within these four regions were spatially averaged for each of the four observation times. The results were compared in time series considering the uncertainty from both in-situ instrumental error and remote sensing temperature retrieving algorithm.

ECR_Final grid system

Region A

Region B

Region C

Bottom Elev (m)
223.8 [Time 1.000] 240.5
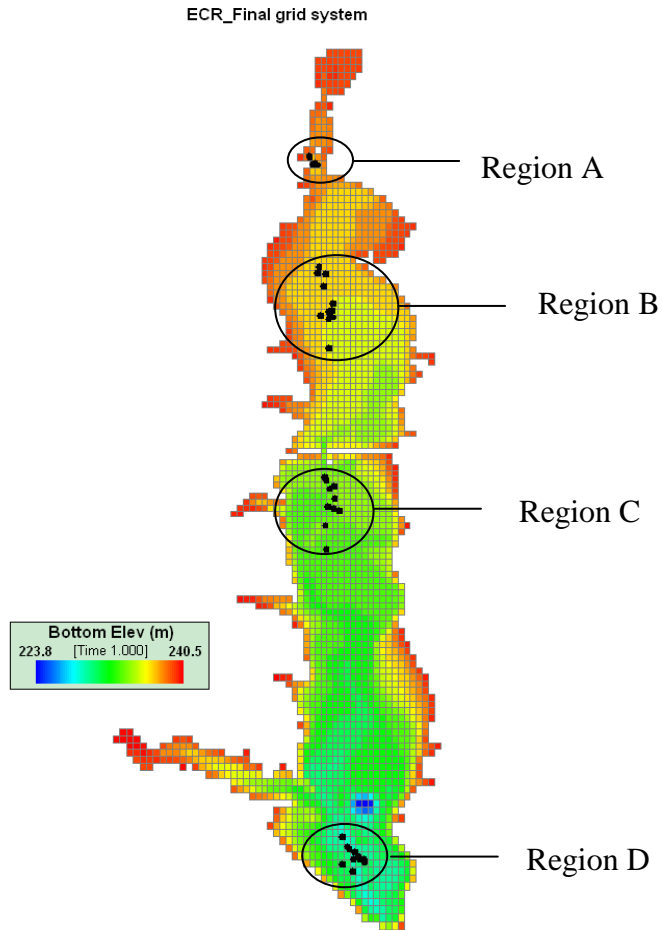
Region D

Figure 15: Four geographical sampling regions in ECR.

In Figure 16 (a-d), four remote sensing observations were interpolated using a third order polynomial, with the trend line very well following the change of observations in time series. New datum was computed with ±1 ⁰C interval for remote sensing data and ±0.15 ⁰C interval for in-situ measurements. When the computed datum was within this interval, the difference was considered to be zero; otherwise, the smaller RMSE of the computed datum was taken. Statistical results indicated the minimum RMSE between remote sensing data and in-situ measurements as 2.47 ⁰C, 3.39 ⁰C, 3.50 ⁰C and 3.95 ⁰C for each remote sensing day respectively. An average of 3.33 ⁰C was computed as the overall discrepancy with CEES observations obviously higher than remote sensing data.

This value should be taken into account in the future work to reduce the discrepancy and to match the accuracy level of two data sets before remote sensing data assimilation is applied.
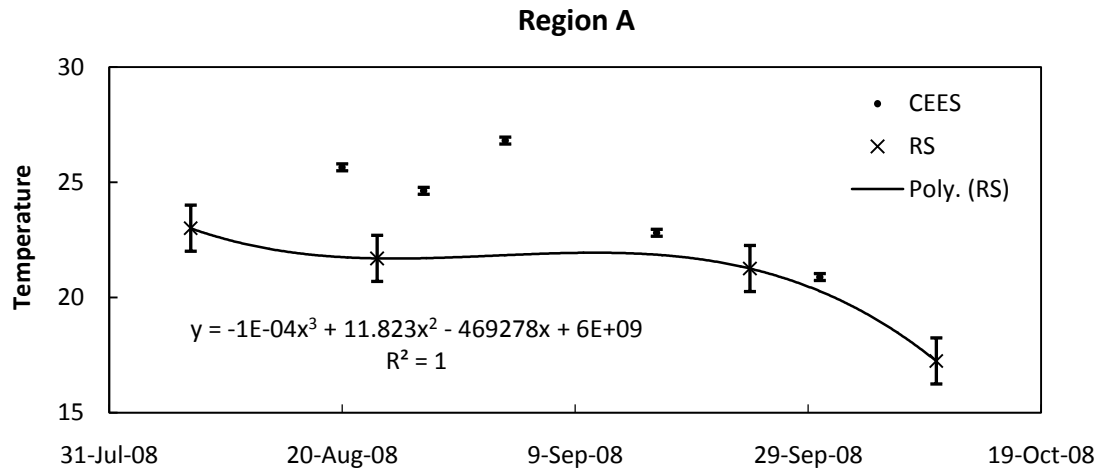
Region A

Figure 16 (a): Comparison between CEES measurements and remotely derived temperature - Region A.
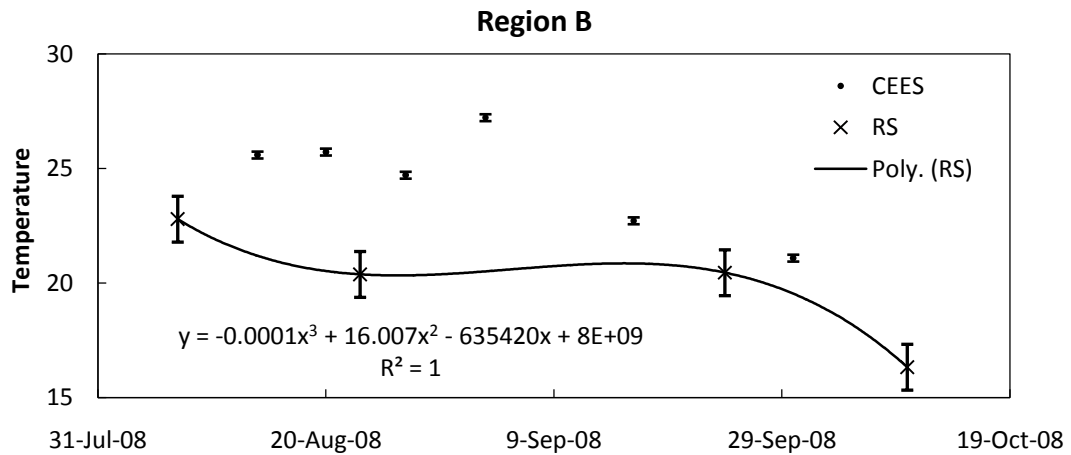
Region B

Figure 16 (b): Comparison between CEES measurements and remotely derived temperature - Region B.

Region C

$$y = -0.0001x^3 + 15.264x^2 - 605941x + 8E+09$$
$$R^2 = 1$$
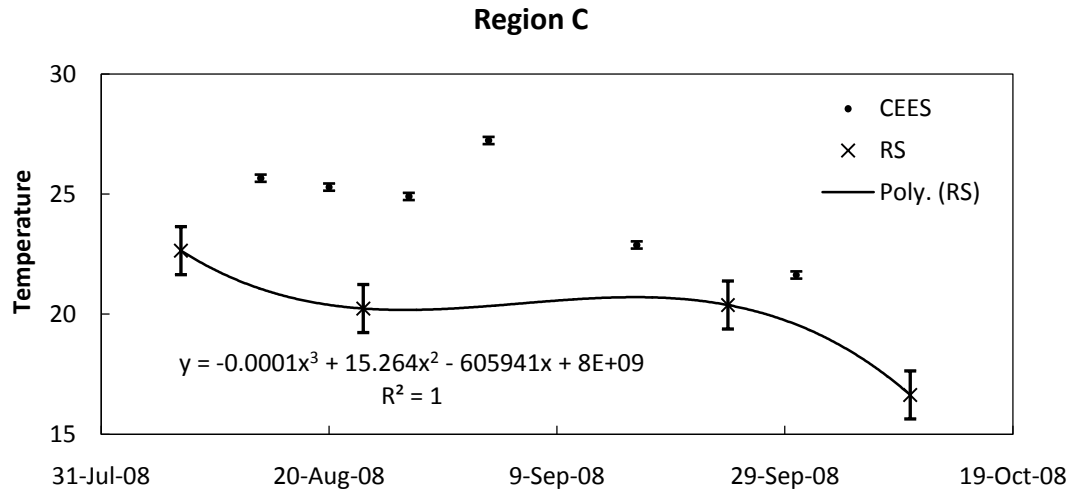
Figure 16 (c): Comparison between CEES measurements and remotely derived temperature - Region C.

Region D

$$y = -0.0001x^3 + 15.354x^2 - 609545x + 8E+09$$
$$R^2 = 1$$
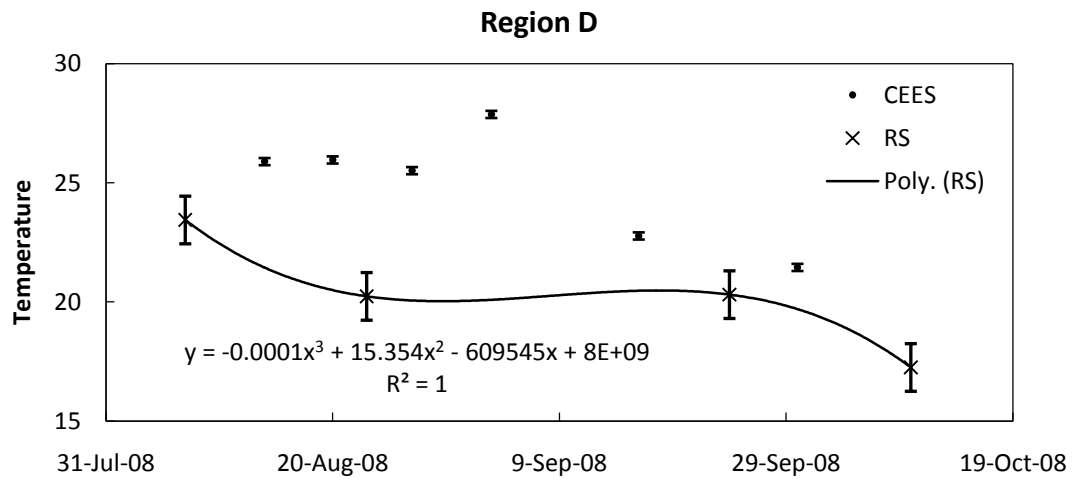
Figure 16 (d): Comparison between CEES measurements and remotely derived temperature - Region D.

4   CONCLUSION AND DISCUSSIONS

This research explored the possible remote sensing data assimilation methods applied in a finite difference hydrodynamic model in order to improve the model capability of simulating real-time variations from multiple data sources. To model the three-dimensional hydrodynamic process within Eagle Creek Reservoir (ECR), Central Indiana, Environmental Fluid Dynamics Code (EFDC) was used. Four remote sensing images obtained from Landsat-5 ETM+ band 6 provided water temperature data, which were assimilated into the EFDC hydrodynamic model. A variational data assimilation approach was applied with continuous adjusting the model initial conditions of the assimilation window. The model was reinitialized and optimized through a genetic algorithm to minimize the difference between the model outputs and the remote sensing observation. The ECR water surface elevation calibration introduced a correction to the watershed model (SWAT) output discharge. The water surface calibration provided an RMSE as 0.029 meter between the model outputs and the USGS measurements. The calibration of water column temperature resulted in an RMSE as 1.279 ⁰C compared with the in-situ measurements. From the model results, EFDC water surface elevation and water column temperature calibration in ECR was fairly proved with model capability of simulating physical parameters within satisfactory bounds of accuracy during the first seven months of 2008.

During the procedure of data assimilation, the genetic algorithm successfully generated model initial conditions which reduce the difference between the model outputs and the remote sensing observations. The final result of data assimilation was determined based on Data Assimilation Efficiency (DAE). The corresponding percentage

38

changes in model initial conditions provided a RRMSE as 0.192 and RMSE as 2.332 °C

between the model outputs and remote sensing retrieved data. After data assimilation, as

the background model was adjusted to provide outputs which are closer to the remote

sensing observations, its outputs had increasing difference with the in-situ measurements.

Therefore, the assumption that the remote sensing data consists better with the field

measurements compared with model outputs was rejected. This result could possibly be

explained from the view of the water depth that remote sensing signal can penetrate into.

With the limitation of EFDC, only evenly distributed vertical layer depth could be

generated. It means that the surface layer outputs from the numerical model represent

properties with water depths which are varying with total depth. However, the remote

sensing technology only measures the water skin temperature. Thus, this comparison

between model outputs and remote sensing data might introduce errors in data

assimilation process. Other than this, every observation has its systematic and random

errors. With the systematic error estimated from both remote sensing retrieving algorithm

and instrumental measurements, the difference between these two sets of data was

computed to estimate the offsets. The results showed an average of 3.33 °C RMSE. In the

future work, the error estimation of observations and conceptual models is going to be

taken into evaluating the data assimilation performance. It is ideal to take both

observation and numerical model errors into the data assimilation procedure.

After data assimilation was applied, it was discovered that the variational data

assimilation algorithm has some limitations in this case study. By defining the

assimilation window as covering all of the four observations and adjusting the initial

conditions, the data assimilation effects were reduced as time propagates in assimilation

window. It was indicated that the initial condition as model adjustment factor had limited effects on model performance with its propagation. Comparatively, the regular input model boundary conditions adjusted the model with boundary inputs and might have more significant influence on the model outputs at every time step. In addition, the variational data assimilation algorithm required significant computational resources considering the finite difference model propagation.

EFDC validation for water surface elevation indicated an RMSE as 0.085 meters between the modeled and USGS measurements. Model validation of water column temperature was carried out for model performance before and after remote sensing data assimilation is applied. The RMSE between model outputs and CEES in-situ measurements for water temperature increased slightly from 0.960 $^{\text{o}}$C to 0.962 $^{\text{o}}$C after data assimilation. However, this value was significantly lower than the model error during calibration and data assimilation period, where the RMSE was 1.279 $^{\text{o}}$C. The validation results indicated the fairly accurate capability of model predictions.

This research addressed the challenge of bringing model outputs, remote sensing data and in-situ measurements into the same spatial-temporal scale. To solve this challenge, a number of assumptions were made. And the possible error sources were explained as follows:

1)   The data assimilation algorithm was performed based on model outputs and remote sensed ETM+ images. Horizontally, it was assumed that each random location was fully representative of the area round within one 30m*30m remote sensing pixel and model grid. Since the spatial resolution is smaller for ETM+ images than for model grid system, the unit pixel for each data set would be mismatched in spatial scale.

40

For example, when two data sets were overlapped with each other, there would be more than one pixel lain over the single model grid. When the random locations were picked up by points, the corresponding value extracted from the ETM+ images would not necessarily be the dominant pixel value located within its model grid size. Similarly, when identifying the in-situ measurement locations, the corresponding values would not necessarily be dominant pixel value located within its model grid size from the ETM+ image. Thus assimilating corresponding remote sensed water temperature into the model grids would introduce errors during the horizontal integrating process.

      2)   Vertically, the remote sensing signal received by the satellite sensor has the dominant information from the water skin temperature, which is less than 0.05 meter. It was assumed that the integrated first layer outputs from the numerical model were fully representative of the skin temperature at water surface. However, due to the limitation of model grid generation, the first layer depths vary according to the total water depth in different regions. Thus assimilating the water skin temperature to adjust the 3-D model might possibly introduce errors.

      3)   When comparing the remote sensing data with in-situ measurements, it was assumed that the measurements at 0.25 meter depth were able to be compared with remote sensed water surface temperature. However, as the remote sensing technology can only represent the top surface of the water column, the difference between the water temperature on the skin and that at 0.25 meter depth was omitted. This difference should be examined according to the temperature gradient at the time of in-situ measurements to avoid the influence cause by water depth.

Proposed method to improve the current research is shown as follows:

1)   To avoid calibrating and improving the model for the first three seasons (spring, summer and fall), and validating for winter season, it is proposed that only three TM images are used for data assimilation instead of four. One remote sensing event and all the in-situ measurements during September 24th and October 10th will be included into the model validation. The model will continuously be validated with in-situ measurements from October 10th to December 31st, 2008.

2)   Since currently there is no proof of the systematic and random human error source from either remote sensing or in-situ measurements, assimilating only remote sensing data with the "truth" assumed to be the in-situ measurements will possibly result in the bias estimation of data assimilation performance. In addition, using only remote sensing data, the numerical model was adjusted according to the surface layer condition. However, the in-situ measurements provide observations from a three-dimensional view of water column parameters. Therefore it is proposed that both remote sensing and in-situ measurements can be assimilated at the same time during the data assimilation period. In this case, a multi-objective Non-dominated Sorting Genetic Algorithm (NSGAII) will be applied as the optimization scheme. The two objectives include the difference between model outputs and remote sensing observations, and the difference between model outputs and in-situ measurements. NSGAII continues to seek from both objectives the combination of individuals, which lie on the Pareto front. Thus the Pareto-optimal solutions will be determined based on the appropriate weights respectively given to the objectives. Meanwhile, a single objective genetic algorithm is going to be applied with only remote sensing data assimilated into the EFDC. The data

assimilation performance will be evaluated with results given by multi-objective NSGAII and single-objective genetic algorithm using the validation data from September 24th to December 31st, 2008.

APPENDICES

Appendix A. Supplementary literature review of EFDC

In recent years, Environmental Fluid Dynamics Code (EFDC) has been widely used for hydrodynamic and water quality modeling. For study cases with complex hydrodynamic processes, a 3-D model can provide precise dynamic simulations of flow, contaminant fate and transport, and biochemical interactions. Multiple applications of EFDC currently exist in the literature. These applications have investigated hydrodynamic and/or water quality simulations in rivers (Ji et al., 2002; U. S. Army Corps of Engineers, 2004b; U. S. Army Corps of Engineers, 2004a; Tetra Tech, 2005; Tetra Tech, 2007a), lakes (USEPA, 2002; Scientifc Environmental Applications, 2003; Jin and Ji, 2004; Zou et al., 2006; Tetra Tech, 2009), reservoirs (Khangaonkar et al., 2005), estuaries (Shen et al., 1999; Tetra Tech, 2001; Wool et al., 2003; Tuckey et al., 2006) and coastal areas (Tetra Tech, 2006; Tetra Tech, 2007b).

Appendix B. Detailed applications of direct and dynamic observer data assimilation methods.

Among direct observer data assimilation algorithms, most applied Kalman Filter related methods including traditional Kalman Filter, Extended Kalman Filter and Ensemble Kalman Filter have been widely used for data assimilations in environmental models. Different application objectives include surface water (Madsen, 2006; Voutilainen et al., 2007), land surface (Launay and Guerif, 2005; Ines et al., 2006; Moradkhani, 2008; Shen et al., 2009), air quality (Vijayaraghavan et al., 2008) and ecosystem (Chemin et al., 2004; Mo et al., 2008; Bouarifi and Alaa, 2010). However, for most of the study cases using direct observer method, regular in-situ measurements were assimilated. A typical case where the extended Kalman Filter is applied aimed to forecast algal bloom dynamics combining with an ecosystem model (Mao et al., 2009). In this study, high frequency field observations (sampling interval $\Delta t = 1$ day, 2h, 1h respectively) and bi-weekly nutrient data are assimilated into the model to produce combined state estimates. (Allen et al., 2003) used an Ensemble Kalman Filter (EnKF) to assimilate long time series of in-situ Chlorophyll taken from a data buoy in the Cretan Sea. The results show a significant improvement in the ecosystem model to forecast chlorophyll concentration.

In recent studies, most of dynamic observer (variational) data assimilation methods have been applied in research of land surface vegetations. (Ines et al., 2006) characterized a typical irrigated agriculture system using a variational remote sensing data assimilation. A modified genetic algorithm was used in data assimilation and water management optimizations. In another study by (Dente et al., 2008), the Leaf Area Index

45

was retrieved from ENVISAT ASAR and MERIS data and was assimilated into CERES-Wheat crop growth model through a variational method. The model was continuously reinitialized with the optimal input parameters to achieve a pixel by pixel agreement between the modeled and remotely sensed observations. Recent research in assimilating Normalized Difference Vegetation Index (NDVI), satellite radar back scattering coefficient (Mangiarotti et al., 2008) also searched for model's initial conditions to minimize their difference to the observed values. The optimization process was realized by introducing an evolutionary strategy. (Boussetta et al., 2008) also proposed a coupled land-atmosphere satellite data assimilation system for a land surface model to provide optimized initial conditions for the atmospheric model.

Beside the land surface parameters, variation data assimilation method was also applied in a number of studies in surface water dynamic processes. Among these studies, in-situ observations were mainly assimilated such as streamflow, precipitation, potential evaporation (Seo, 2001) to explore the hydrologic forecasting and velocities for simulating sea circulation (Panteleev et al., 2007). An adjoined model of a finite-element shallow water equations model was involved to generate the corresponding observations, which were available at each time step of the assimilation window. The cost function was the weighted sum of the difference between the model solution and the generated observations. Successful minimization of the cost function was obtained using initial conditions as control variables (Zhu et al., 1994). A limited memory quasi-Newton method was used as the optimization algorithm to obtain the minimum error between modeled and observed river sediment transports. The control variable was chosen as the initial river discharge (Yang and LeDimet, 1998).

Appendix C. Finite difference grid system generation process

Since EFDC allows single focal point with expanding Cartesian grids, three grid systems were initially generated with the focal point at the water intake to better simulates the complex flow around the gates and near the causeway. Additionally, grids were rotated parallel with the dam in regards of the accuracy of model's capability of simulating flow patterns near the dam. In the initial trial runs, the grid size for the fine grid system was 15.5m with the total 5290 grids through ECR, for the medium grid system 20.0m with a total 3821, and for the coarse system 35.0m with a total amount of 1763 (Appendix Table 1) (Appendix Figure 1). However, considering the computational costs for the following water quality model, coarser grids which significantly reduce the cell amount and the required time steps were required. Meanwhile, cell rotation angle was also removed in order to make sure water flows through the 56th bridge without any unrealistic physical obstacle. Vertically, EFDC only allows standard sigma vertical grid system, in which the same number of vertical layers was set up throughout the research domain. The number of vertical layers was set to be five, taking account of the photic depth in the deepest area of the reservoir and the computational burden of the numerical model.

Appendix Table 1: Parameter selection for grid generation.

Focal Point (X, Y)	Number of grids	Minimum cell size (m)	Maximum cell size (m)	Expanding factor	Cell rotation angle (degree)
559313,4411545	5290	15.5	100.0	1.020	-52
559313,4411545	3821	20.0	100.0	1.020	-52
559313,4411545	1763	35.0	80.0	1.020	-52
559313,4411545	2401	40.0	60.0	1.005	0
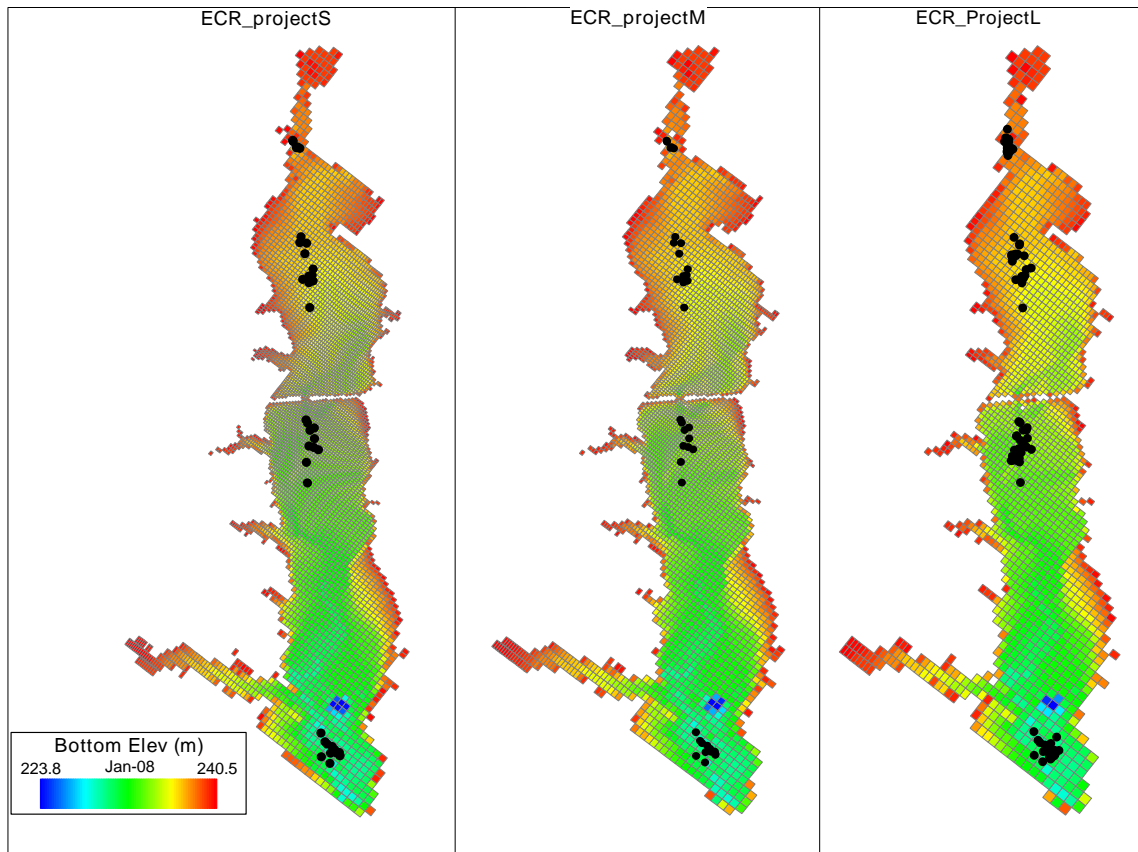
Appendix Figure 1: 2008 water quality sampling locations within three grid systems of in ECR. (Left: fine grid; Middle: medium grid; Right: coarse grid)

Appendix D. Pseudo-code of genetic algorithm application to EFDC model

---

parallel for each individual p in population $\mathcal{P}$

    Initialize(p)

    CallEFDC(p)   // run EFDC with initially set individuals

    EvaluateRRMSE(p)   // costfunction to measure the difference between model and the observations

end parallel for

while $t\ <\ number\ of\ generation$

    $\mathcal{C} = MakeNewPopulation(\mathcal{P})$

    //use selection, crossover and mutation

    parallel for each individual p in new populaiton $\mathcal{C}$

        CallEFDC(p)

        EvaluateRRMSE(p)

    end parallel for

    $\mathcal{R} = \mathcal{P} \cup \mathcal{C}$ //combine $\mathcal{P}$ and $\mathcal{C}$

    $\mathcal{R} = Sort(\mathcal{R})$ // sort in ascending order

    $P = SelectNewPopulaiton(\mathcal{R})$ // based on the minimum cost function results in the previous generation

    $t = t + 1$

end while

Appendix E. EFDC simulations in water surface elevation and mass balance using uncalibrated SWAT

Initial simulations using EFDC were performed for one vertical layer to ensure model stability. Atmospheric boundary conditions were further specified in full hydrodynamic module with multiple layers. After initial hydrodynamic trial runs, five vertical layers were defined for the water column based on computational accuracy as well as time consumption.

With the preliminary results from the on-going Eagle Creek watershed model, a series of EFDC uncalibrated runs were carried out before model calibration. Problems such as negative depth and exceeded iterations would occur because of the inaccuracy of bathymetry data for cells around, impropriate choice of dry depth and numerical solutions. The purpose of this process was to warm up the model, determine proper values for numerical parameters (finite difference model scheme, numerical solution, etc.) which influenced the model iteration, and achieved a basic idea of the model's capability of capturing the trend of hydrodynamic processes.

Based on water balance calculations for ECR for 2008, ECR received total of 2096.94 m^3/s from twelve different tributaries where flow discharge was estimated using an uncalibrated watershed model (SWAT), as well as precipitation (total 1.37 m/year) in 2008. With data collected by water intake and dam operators, the total amount of water out from ECR is estimated to be 2574.905 m^3/s. An average evaporation was taken with the value of approximately 5.5 mm/day from June to October, and an average of 4.01 mm/day from November to May as presented in the previous research (Lobligeois, 2009) in the same domain. Appendix Figure 2 shows the total inflow from SWAT model versus

the total outflow from dam, water intake and evaporation in 2008. According to the calculation, there were +28,250,531.93 m^3 net storage of water in the reservoir in 2008. Positive storage came from the uncalibrated inflow data from the tributaries. In the following work, the inflow and outflow of water in ECR are further considered. A correction step was highly required.

ECR inflows and outflows in 2008
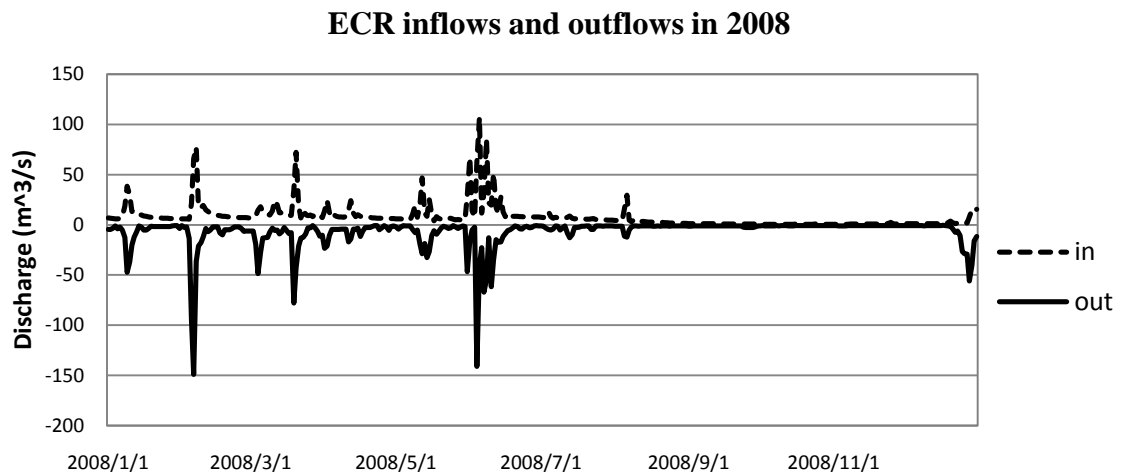
Appendix Figure 2: ECR inflows and outflows in 2008.

Meanwhile, a comparison was done between USGS water surface elevation measurements and model predicted surface elevation based on volume – elevation relationship provided by EFDC (Appendix Figure 3). This relationship was based on topographic data for the reservoir and the modeling grid system that were generated.
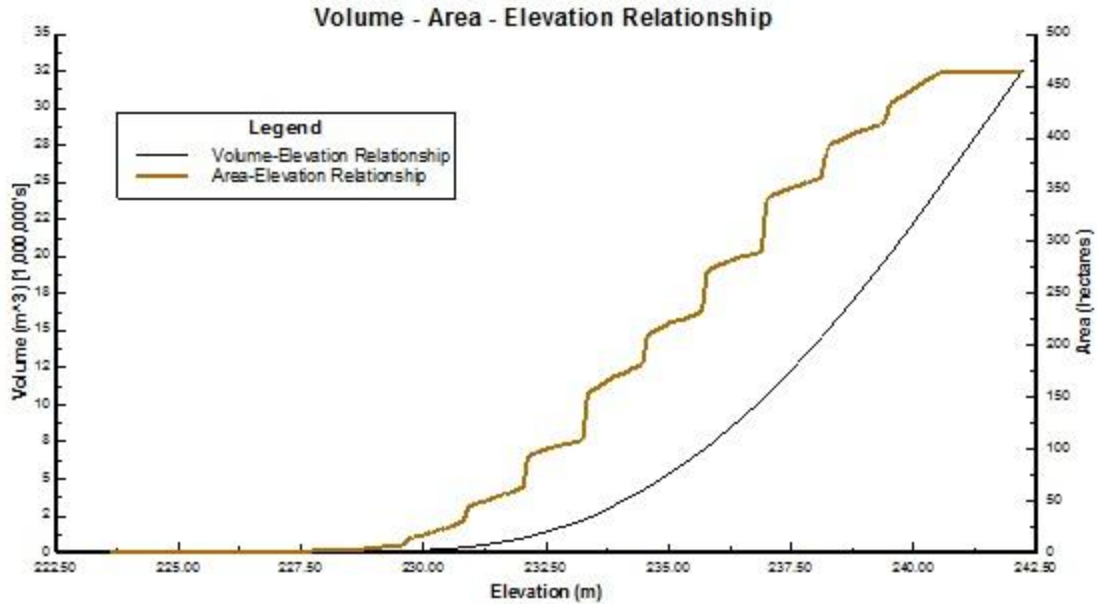
Appendix Figure 3: Volume, area and surface elevation relationship provided in EFDC.

In order to estimate pool elevation based on the relationship between volume and surface elevation of ECR, the net storage of the reservoir was calculated daily in 2008. Discrete data from the curve are imported into Matlab to generate a proper fitting curve and the corresponding functions.

The initial total water volume in the reservoir was calculated using the mathematical relationship from the functions with the known initial water surface elevation. The water surface elevations for the rest of days in 2008 were calculated using the algebraic summation of the water volume within the reservoir in the previous day, and the net storage of water volume that was calculated from the inflow and outflow in that day. In hydrodynamic calibration of EFDC, water surface elevation for ECR would be compared with in-situ data measured near the dam from USGS database. Appendix Figure 4 shows the simulated ECR surface elevation in the first uncalibrated run of EFDC for 2008. The results were compared with mass-balanced water surface elevation from

uncalibrated SWAT flow, and USGS gauge station measurements to the east of Eagle

Creek dam.
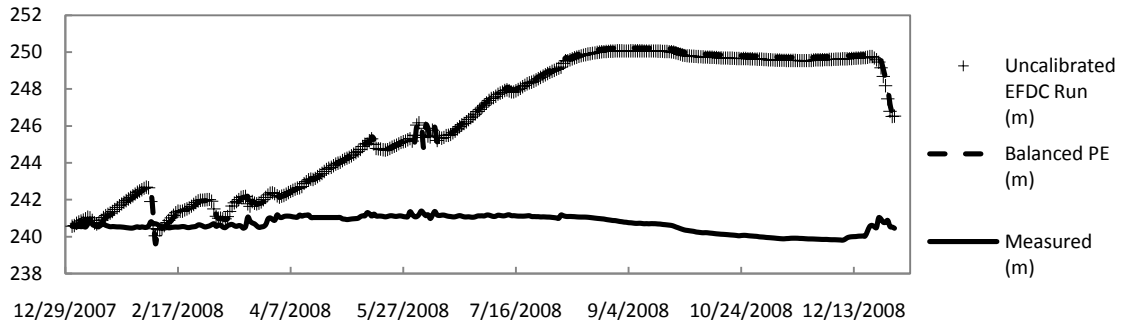
Appendix Figure 4: Modeled pool elevation using uncalibrated SWAT inflows (crosses), measured pool elevation near the dam (dashed line), and Volume-Elevation based pool elevation calculated from uncalibrated SWAT inflows (continuous line).

Although there existed an obvious difference between modeled and in-situ data, the results of the comparison (Appendix Figure 5) showed excellent agreement ($R^2 =$ 0.99) between simulated and calculated pool elevation from water balance. Since EFDC reads input flows from each tributary provided by SWAT model outputs, the SWAT flow data are the only 'knowledge source' for the model. Given that its computation performs in good correlation with input knowledge, it was concluded that the uncalibrated flow data from watershed model is a significant contributing factor to the discrepancies in water surface elevation. Therefore, in the calibration of EFDC, further calibration of SWAT model was required.
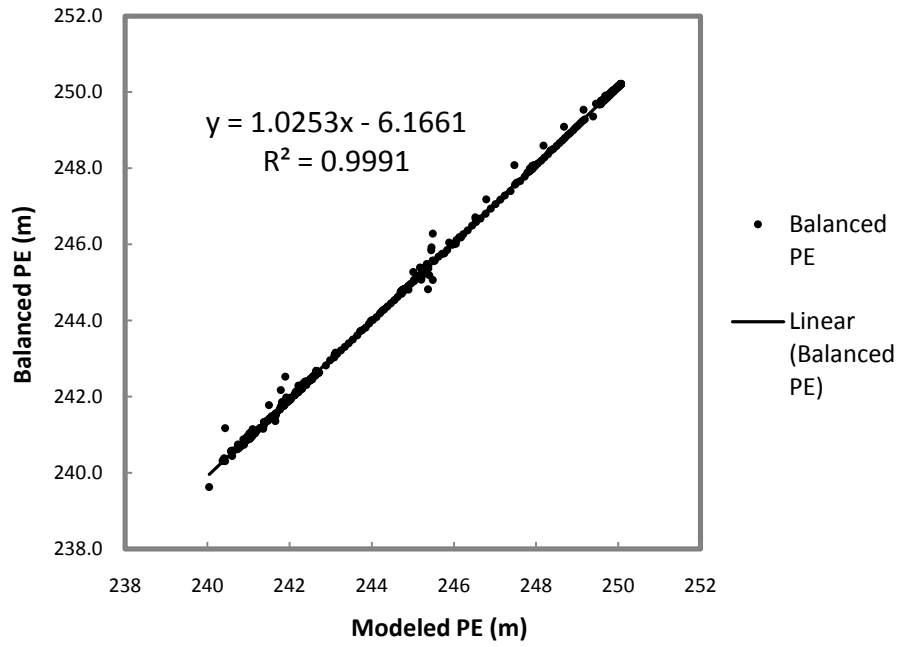
Appendix Figure 5: Comparison between modeled and calculated pool elevation.

Appendix F. Initial and boundary condition for water temperature

Water temperature in January 2008 was required for initial condition of the model. However, two watershed monitoring stations established by CEES (ECWMP-01 and ECWMP-02) did not collect daily water temperature and could not be used as data sources (Appendix Figure 6). Meanwhile, water temperature for ECR was not collected during winter months and data from buoy placed in reservoir were not available for this period. Thus, data from nearest USGS monitoring stations was used (Appendix Figure 7) for initial and boundary conditions. Two stations provide most accurate water temperature for 2008 are USGS gauge Station #03359000 at Mill Creek (located in a different eco-region) and USGS gauge Station #03354000 at White River near Centerton, IN (approximately 25 miles south from ECR). Water temperature for tributaries from January 1 to January 13, 2008 was determined based on a linear relationship between USGS Centerton station and Mill Creek station. USGS Stout Gen. Station #03353611 which was closer to the study domain could not be used because it was reported to be thermally polluted. Water temperature data for boundary conditions for the rest of year 2008 came from USGS Centerton station (Appendix Table 2).

Appendix Table 2: Boundary water temperature data sources for 2008.

	ECWMP-01	ECWMP-02	USGS Stout Gen. Station # 03353611	USGS Centerton Station # 03354000	USGS Mill Creek Station # 03359000
Data Provided	6 days	4 days	Jan 14th – Dec 31st	Jan 14th – Dec 31st	Jan 1st – Dec 31st
Data Description	Inadequate data source		Thermally polluted	Located in lower White River basin	Located in a different eco-region

ECWMP-02
558,257.76, 4,415,341.54

**ECWMP-01**
557,520.61, 4,409,799.28
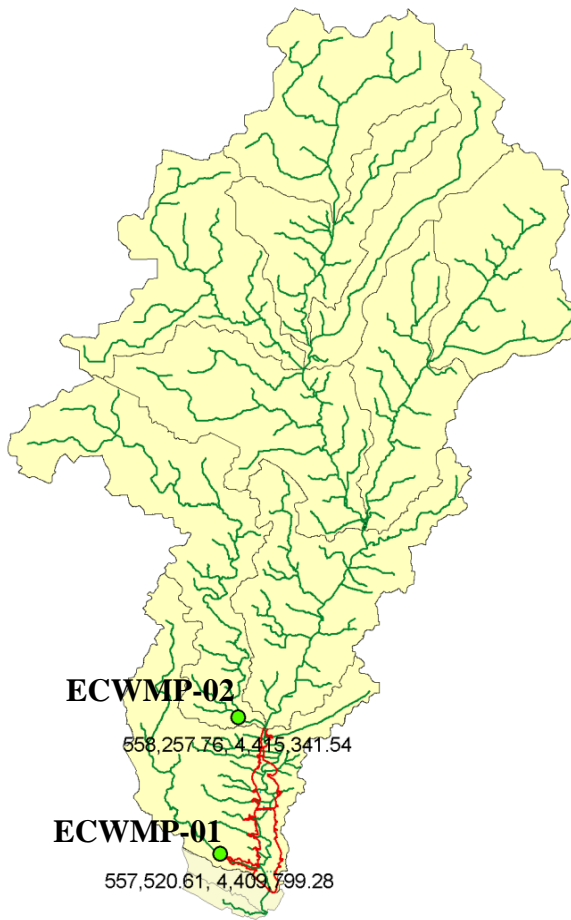
Appendix Figure 6: Eagle Creek watershed water temperature sampling locations close to ECR.

Appendix Figure 7: USGS gauge station locations used in water temperature comparison.
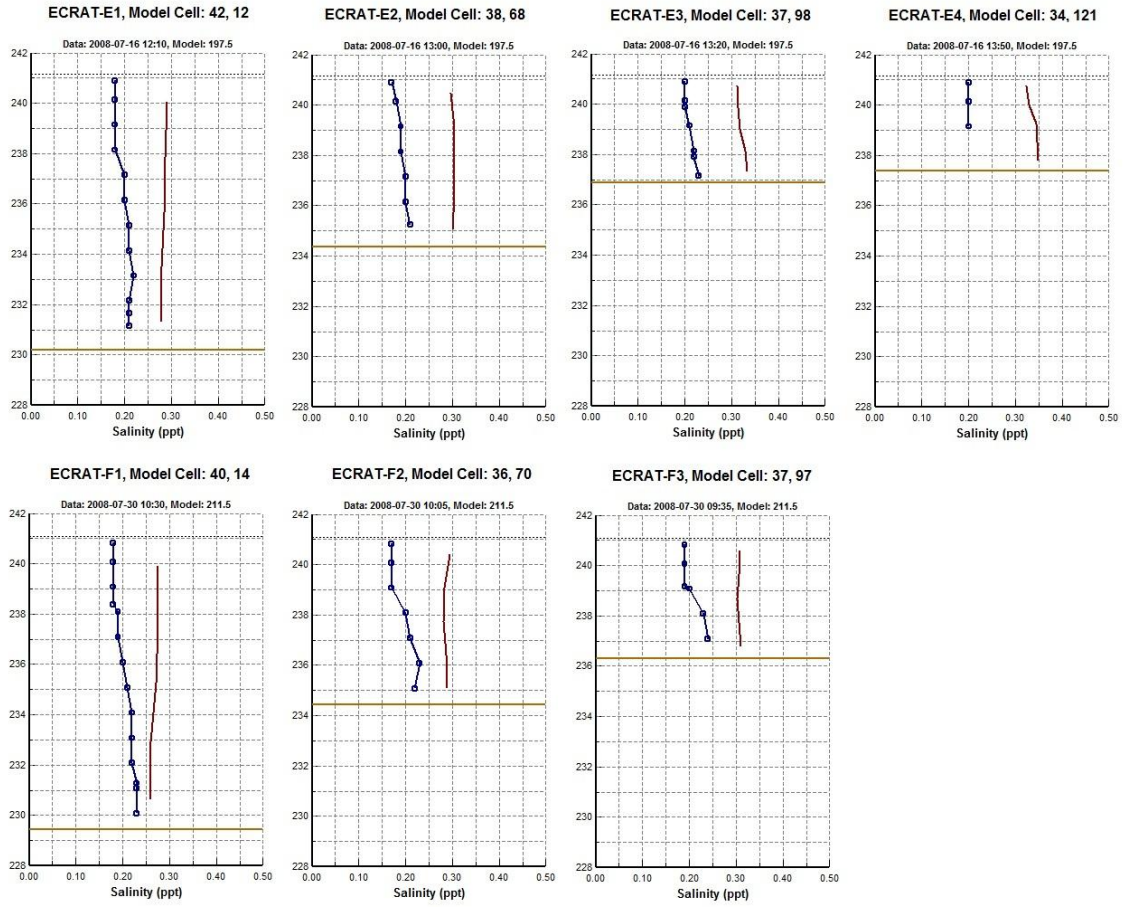
Appendix G. Water column salinity calibration

The in-situ measurements of water column salinity were taken along with water temperature sampling using YSI probes. Initial salinity on January 1st, 2008 was determined as 0.25 kg/m^3, which was the average observation value for ECR in 2008. Salinity calibration was implemented by adjusting the evaporation transfer coefficient in EFDC. The result showed less satisfactory correspondence at this location all through the calibration process, achieving an RMSE as much as 0.100ppt in the results (Appendix Figure 8 (a-c)) (Appendix Table 3). However, the measured salinity in this inland reservoir system was within the range of 0.160 ppt – 0.310 ppt, and the modeled range as 0.268 ppt – 0.350 ppt. Previous research (Somchai Wangwibulkit, 2008) showed very limited influences on the concentration of Chlorophyll-a with the water column salinity less than 2.00 ppt. Both of the model calibrated and the measured salinity for this inland water system were satisfactorily within this range, which indicated that salinity value for this study case was not a dominant factor for the simulation and prediction of algal bloom events.
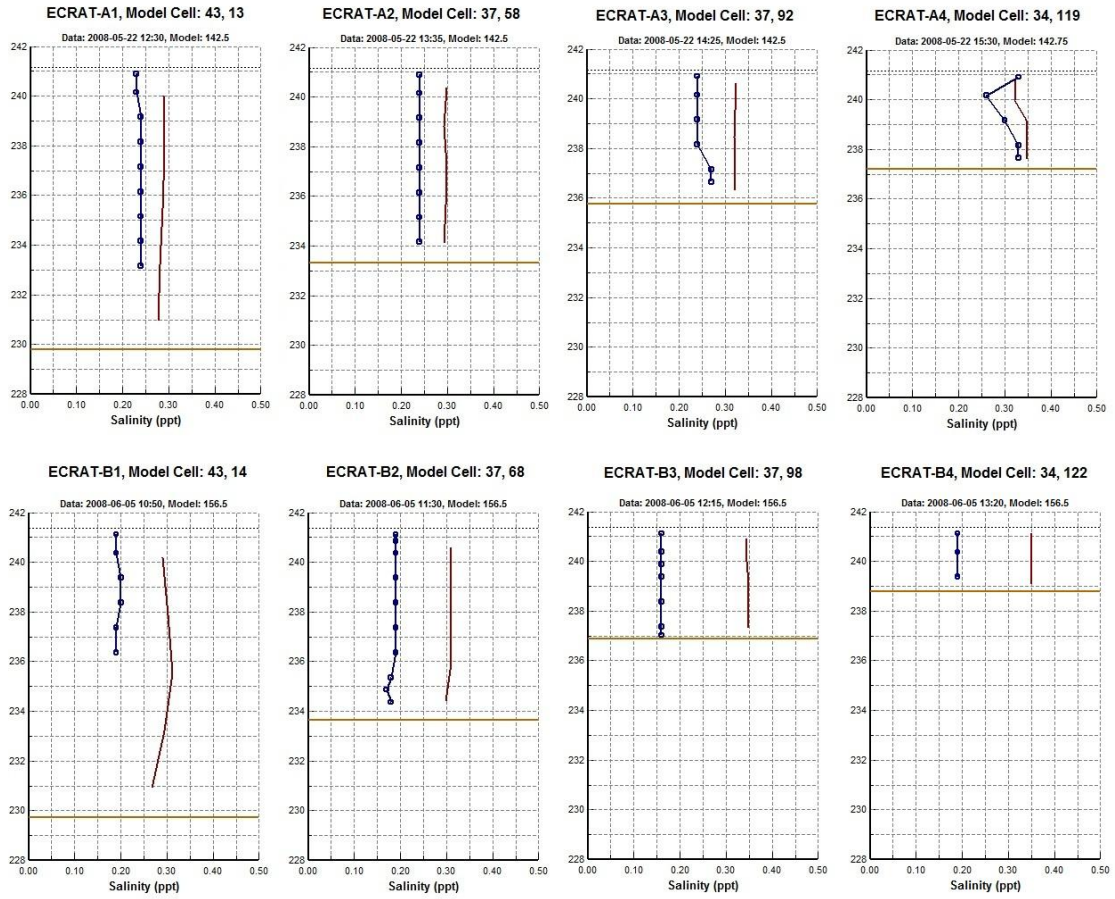
Appendix Figure 8 (a): Calibrated salinity vertical profiles for ECR.
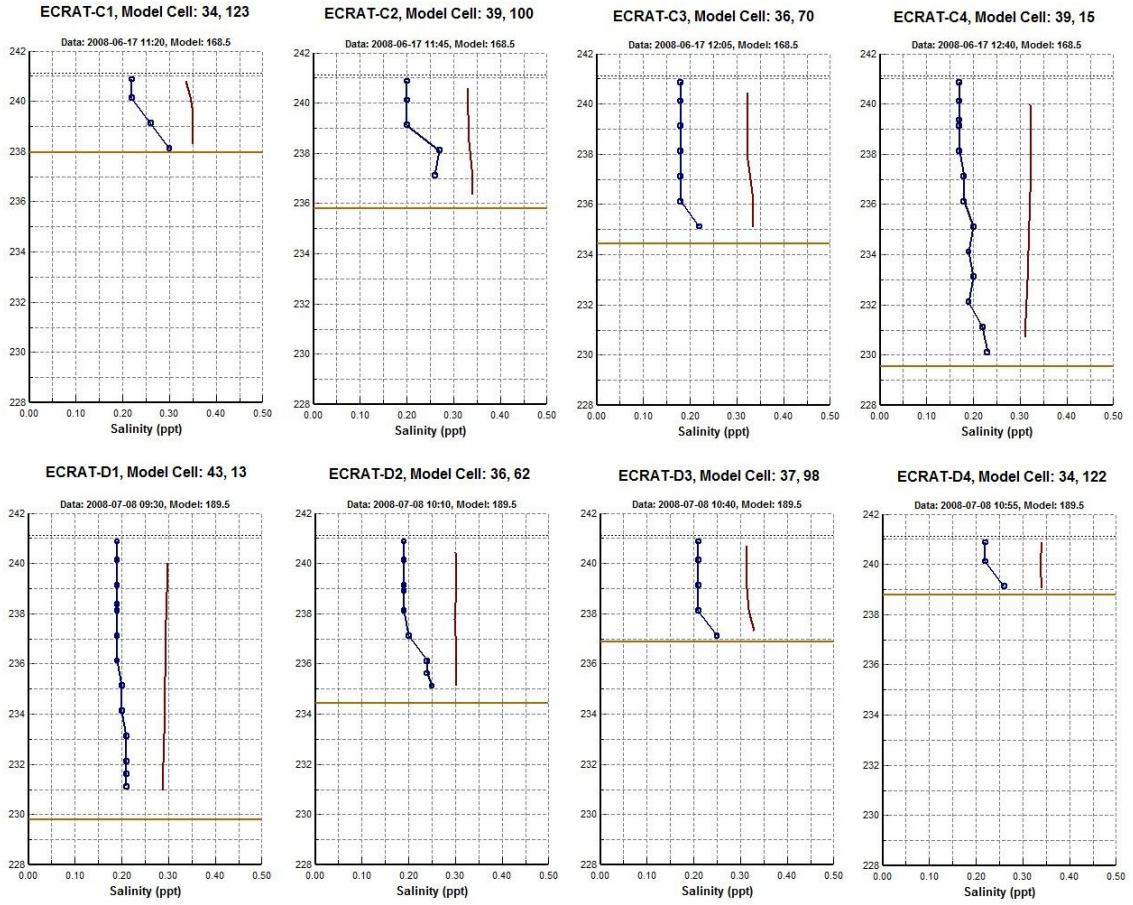
Appendix Figure 8 (b): Calibrated salinity vertical profiles for ECR.

Appendix Figure 8 (c): Calibrated salinity vertical profiles for ECR.

Appendix Table 3: Statistical review of water column salinity calibration results.

Station ID	Date/Time	# Pairs	Parameter	RMSE (ppt)
ECRAT-A1	22-May-08	9	Salinity	0.050
ECRAT-A2	22-May-08	8	Salinity	0.056
ECRAT-A3	22-May-08	6	Salinity	0.073
ECRAT-A4	22-May-08	5	Salinity	0.037
ECRAT-B1	5-Jun-08	11	Salinity	0.105
ECRAT-B2	5-Jun-08	10	Salinity	0.121
ECRAT-B3	5-Jun-08	7	Salinity	0.187
ECRAT-B4	5-Jun-08	5	Salinity	0.160
ECRAT-C1	17-Jun-08	5	Salinity	0.092
ECRAT-C2	17-Jun-08	5	Salinity	0.112
ECRAT-C3	17-Jun-08	8	Salinity	0.137
ECRAT-C4	17-Jun-08	13	Salinity	0.133
ECRAT-D1	8-Jul-08	13	Salinity	0.096
ECRAT-D2	8-Jul-08	9	Salinity	0.095
ECRAT-D3	8-Jul-08	5	Salinity	0.100
ECRAT-D4	8-Jul-08	4	Salinity	0.095
ECRAT-E1	16-Jul-08	12	Salinity	0.086
ECRAT-E2	16-Jul-08	7	Salinity	0.111
ECRAT-E3	16-Jul-08	7	Salinity	0.110
ECRAT-E4	16-Jul-08	3	Salinity	0.132
ECRAT-F1	30-Jul-08	14	Salinity	0.069
ECRAT-F2	30-Jul-08	9	Salinity	0.092
ECRAT-F3	30-Jul-08	6	Salinity	0.102
Composite Statistics		185 (sum)	Salinity	0.100 (average)

REFERENCES

ADEM/Water Quality Branch, 2008. Final Coosa River Basin Total Maximum Daily
   Loads for Nutrients, OE/DO and PH, Alabama Department of Environmental
   Management.

Allen, Eknes and Evensen, 2003. An Ensemble Kalman Filter with a complex marine
   ecosystem model: hindcasting phytoplankton in the Cretan Sea. *Annales
   Geophysicae* **21**(1): 399-411.

APEC (2010). "Can drinking water reservoirs develop manganese problems due to
   temperature stratification?". from
   http://freedrinkingwater.com/water_quality/quality1/40-08-lake-stratification-and-
   manganese.htm.

Bouarifi and Alaa, 2010. A Global Optimization Approach to Assimilate the Initial
   Conditions into a SVAT Model. *International Journal of Geoinformatics* **3**(1):
   42-55.

Boussetta, Kolke, Yang, Graf and Pathmathevan, 2008. Development of a coupled land-
   atmosphere satellite data assimilation system for improved local atmospheric
   simulations. *Remote Sensing of Environment* **112**(3): 720-734.

Caliskan, 2008. Modeling of Hydrodynamics and Sedimentation in a Stratified Reservoir:
   Thahtali Reservoir, Izmir. Department of Civil Engineering. Izmir, Izmir Institute
   of Technology. **Master**.

Chemin, Honda and Ines, 2004. Genetic algorithm for assimilating remotely sensed
   evapotranspiration data using a soil-water-atmosphere-plant model -
   Implementation issues. International Journal of Geoinformatics. **1:** 87-90.

Dente, Satalino, Mattia and Rinaldi, 2008. Assimilation of leaf area index derived from
   ASAR and MERIS data into CERES-Wheat model to map wheat yield. *Remote
   Sensing of Environment* **112**(4): 1395-1407.

Goldberg (1989). <u>Genetic algorithms in search and optimization and machine learning</u>,
   Addison-Wesley Publ. Co. Inc. USA.

Goodin, 1995. Mapping the surface radiation budget and net radiation in a sand hills
   wetland using a combined modeling/remote sensing method and Landsat thematic
   Mapper Imagery. Geocarto International, Taylor & Francis. **10:** 19-29.

Hamerick, 1992. A Three-Dimensional Environmental Fluid Dynamics Computer Code: Theoretical And Computational Aspects, School of Marine Science, Virginia Institute of Marine Science

Houser, Shuttleworth, Famiglietti, Gupta, Syed and Goodrich, 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research* **34**(12): 3405-3420.

Ines, Honda, Das Gupta, Droogers and Clemente, 2006. Combining remote sensing-simulation modeling and genetic algorithm optimization to explore water management options in irrigated agriculture. *Agricultural Water Management* **83**(3): 221-232.

Ji, Hamrick and Pagenkopf, 2002. Sediment and metals modeling in shallow river. *Journal of Environmental Engineering-Asce* **128**(2): 105-119.

Jin and Ji, 2004. Case study: Modeling of sediment transport and wind-wave impact in Lake Okeechobee. *Journal of Hydraulic Engineering-Asce* **130**(11): 1055-1067.

Khangaonkar, Yang, DeGasperi and Marshall, 2005. Modeling hydrothermal response of a reservoir to modifications at a high-head dam. *Water International* **30**(3): 378-388.

Korotaev, Huot, Le Dimet, Herlin, Stanichny, Solovyev and Wu, 2008. Retrieving ocean surface current by 4-D variational assimilation of sea surface temperature images. *Remote Sensing of Environment* **112**(4): 1464-1475.

Launay and Guerif, 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. *Agriculture Ecosystems & Environment* **111**(1-4): 321-339.

Li, Koike and Pathmathevan, 2004. A very fast simulated re-annealing (VFSA) approach for land data assimilation. *Computers & Geosciences* **30**(3): 239-248.

Lobligeois, 2009. Development of a 3D Hydrodynamic Model for Eagle Creek Reservoir.

Madsen, 2006. Water Quality Surveillance and Early Warning in Surface Waters – Intergration of Mathematical Models and On-line Monitoring. 7th International Conference on Hydroinformatics, HIC 2006. Nice, France.

Mangiarotti, Mazzega, Jarlan, Mougin, Baup and Demarty, 2008. Evolutionary bi-objective optimization of a semi-arid vegetation dynamics model with NDVI and sigma(0) satellite data. *Remote Sensing of Environment* **112**(4): 1365-1380.

Mao, Lee and Choi, 2009. The extended Kalman filter for forecast of algal bloom dynamics. *Water Research* **43**(17): 4214-4224.

Mo, Chen, Ju and Black, 2008. Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter. *Ecological Modelling* **217**(1-2): 157-173.

Moradkhani, 2008. Hydrologic remote sensing and land surface data assimilation. *Sensors* **8**(5): 2986-3004.

NASA (2009). Landsat 7 Science Data Users Handbook**:** 106-128.

Olioso, Chauki, Courault and Wigneron, 1999. Estimation of evapotranspiration and photosynthesis by assimilation of remote sensing data into SVAT models. *Remote Sensing of Environment* **68**(3): 341-356.

Panteleev, Proshutinsky, Kulakov, Nechaev and Maslowski, 2007. Investigation of the summer Kara Sea circulation employing a variational data assimilation technique. *Journal of Geophysical Research-Oceans* **112**(C4): -.

Quattrochi (2004). Estimating spatially distributed surface fluxes in a semi-arid Great Basin desert using Landsat TM thermal data. <u>Thermal Remote Sensing in Land Surface Processing</u>. Luvall, CRC Press**:** 133-159.

Rizzoli and Young, 1995. Assimilation of global versus local data sets into a regional model of the Gulf Stream system .1. Data effectiveness. *Journal of Geophysical Research-Oceans* **100**(C12): 24773-24796.

Robinson and Lermusiaux, 2000. Overview of data assimilation. Harvard Reports in Physical/Interdisciplinary Ocean Science.

Scientifc Environmental Applications, 2003. Calibration, Verification, Validation Report. Salinity Distribution and Flow Management Studies for Lake Worth Lagoon.

Seo, Koren and Cajina, 2003. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. *Journal of Hydrometeorology* **4**(3): 627-641.

Shen, Boon and Kuo, 1999. A modeling study of a tidal intrusion front and its impact on larval dispersion in the James River estuary, Virginia. *Estuaries* **22**(3A): 681-692.

Shen, Yang, Li, Tan, Li and Le Toan, 2009. A scheme for regional rice yield estimation using ENVISAT ASAR data. *Science in China Series D-Earth Sciences* **52**(8): 1183-1194.

Slater and Clark, 2006. Snow data assimilation via an ensemble Kalman filter. *Journal of Hydrometeorology* **7**(3): 478-493.

Sobrino, Jimenez-Munoz and Paolini, 2004. Land surface temperature retrieval from LANDSAT TM 5. *Remote Sensing of Environment* **90**(4): 434-440.

Somchai Wangwibulkit, 2008. Effects of Salinity and pH on the Growth of Blue-Green Algae, Oscillatoria sp. and Mocrocystis sp., Isolated from Pacofoc White Shrimp (Litopenaeus vannamei) Ponds. *Kasetsart University Fisheries Research Bullet* **No. 32(1)**: 1-9.

Tetra Tech, 2001. The Lower Cape Fear River Estuary Modeling Report.

Tetra Tech, 2005. A Hydrodynamic and Water Quality Model for the Lower Charles River Basin, Massachusetts.

Tetra Tech, 2006. Development of a Florida Bay and Florida Keys Hydrodynamic and Water Quality Model: Hydrodynamic Model Calibration.

Tetra Tech, 2007a. Green River Sediment Stability Analysis.

Tetra Tech, 2007b. Three-Dimensional Model Development for the Charleston Harbor System.

Tetra Tech, 2009. Hydrodynamic and Water Quality Modeling Report for Lake Lanier, Georgia.

Tuckey, Gibbs, Knight and Gillespie, 2006. Tidal circulation in Tasman and Golden Bays: implications for river plume behaviour. *New Zealand Journal of Marine and Freshwater Research* **40**(2): 305-324.

U. S. Army Corps of Engineers, 2004a. Model Calibration: Modeling Study of PCB Contamination in the Housatonic River. **1**.

U. S. Army Corps of Engineers, 2004b. Three-Dimensional Eutrophication Model of the Lower St. Johns River, Florida.

USEPA, 1999. Uncovered Finshed Water Reservoirs. EPA Guidance Manual. **EPA 815-R-99-011**.

USEPA, 2002. Total Maximum Daily Load for Dissolved Oxygen and Nutrients to Mashapaug Pond, Rhode Island.

USEPA (2010). "Environmental Fluid Dynamics Code (EFDC)." from http://www.epa.gov/ceampubl/swater/efdc/.

Vijayaraghavan, Snell and Seigneur, 2008. Practical Aspects of Using Satellite Data in Air Quality Modeling. *Environmental Science & Technology* **42**(22): 8187-8192.

Voutilainen, Pyhalahti, Kallio, Pulliainen, Haario and Kaipio, 2007. A filtering approach for estimating lake water quality from remote sensing data. *International Journal of Applied Earth Observation and Geoinformation* **9**(1): 50-64.

Walker (2005). Hydrologic Data Assimilation. <u>Advances in Water science Methodologies</u>**:** 1-30.

Wool, Davie and Rodriguez, 2003. Development of three-dimensional hydrodynamic and water quality models to support total maximum daily load decision process for the Neuse River Estuary, North Carolina. *Journal of Water Resources Planning and Management-Asce* **129**(4): 295-306.

Yang and LeDimet, 1998. Variational data assimilation in the transport of sediment in river. *Science in China Series D-Earth Sciences* **41**(5): 473-485.

YSI (2011). "YSI 600XL and 600XLM Specifications." from http://www.ysi.com/productsdetail.php?600XL-and-600XLM-7.

Zhu, Chen, Qin, Li and Wang, 2009. Optimization of ecosystem model parameters using spatio-temporal soil moisture information. *Ecological Modelling* **220**(18): 2121-2136.

Zhu, Navon and Zou, 1994. Variational Data Assimilation with a Variable Resolution Finite-Element Shallow-Water Equations Model. *Monthly Weather Review* **122**(5): 946-965.

Zou, Carter, Shoemaker, Parker and Henry, 2006. Integrated hydrodynamic and water quality modeling system to support nutrient total maximum daily load development for Wissahickon Creek, Pennsylvania. *Journal of Environmental Engineering-Asce* **132**(4): 555-566.

CURRICULUM VITAE

Shuangshuang Xie

EDUCATION

M.S.: Earth Sciences, Indiana University, 2011
B.S.: Geography, Xiaozhuang University, 2009

CAREER/WORK EXPERIENCE

Department of Earth Sciences, Indiana University, Indianapolis, IN
2009 to 2011

Research Assistant: Construct and apply hydrodynamic model using Environmental Fluid Dynamics Code (EFDC)

   Case study: Eagle Creek Reservoir, central Indiana (Cooperated with Veolia Water, Indianapolis)
- Collect field measurements, USGS measurements, and atmospheric data for model construction
- Calibrate hydrodynamic model with calibrated watershed model (SWAT) inflow discharge
- Analyze data of water surface elevation, water temperature and salinity vertical profiles
- Cooperate with water quality model calibration
- Assimilate remote sensing data into the numerical model

School of Geographical Sciences, Xiaozhuang University, P. R. China
2008 to 2009

Research Assistant: Researched on calculation of tourist environmental carrying capacity with ecological footprint model for City of Nanjing
- Collected geographical data in population, area, and economic growth rates from national statistical bureau
- Constructed six sub-models based on the empirical model for calculation
- Data analysis to estimate the fossil energy consumption compared with regional and global supplement

Nanjing Environmental Monitoring Center, P. R. China
2008

Internship: Took part in the project cooperating with the 3rd Investigation of Pollution Headstream around Jiangsu Province
- Conducted in-situ measurements, and collected water samples at study sites
- Cooperated with the laboratory chemical analysis of water samples
- Technical reports writing

School of Geographical Sciences, Xiaozhuang University, P. R. China
2005 to 2007

<u>Research Assistant</u>: Conducted water sample analysis in department laboratory to determine surface water contaminants on campus.

- Determined the sampling locations and collected surface water samples at a campus river
- Analyze water samples using spectrophotometer, Flame Atomic Absorption Spectrometry (FAAS), Gas Chromatography (GC) and Liquid Chromatography (Hamerick) to determine the electro-sensitive contamination components, and to investigate organic pollutants in laboratory

PRESENTATIONS AND PUBLICATIONS

Xie, S., M. Babbar-Sebens, and S. Bruder, "Remote Sensing Data Assimilation in Water Quality Numerical Models for Simulation of Algal Bloom Dynamics," American Society of Civil Engineers (ASCE) Environmental & Water Resources Institute (EWRI) World Water & Environmental Resources Congress 2011 & Related Symposia, Palm Springs, CA, 2011.

Bruder, S., M. Babbar-Sebens, and **S. Xie**, "Framework for Prediction of Spatial-Temporal Distribution of Algal Metabolites in Algal-Bloom Affected Water Bodies," American Society of Civil Engineers (ASCE) Environmental & Water Resources Institute (EWRI) World Water & Environmental Resources Congress 2011 & Related Symposia, Palm Springs, CA, 2011.

Babbar-Sebens, M., **S. Xie**, and S. Bruder, "Development of Eagle Creek Reservoir Water Quality Model based on Environmental Fluid Dynamics Code," Central Indiana Water Resources Partnership Science Meeting, Indianapolis, IN, May 10, 2010.