  

**Project Eddie: MODELING CLIMATE CHANGE EFFECTS ON LAKES**

**USING DISTRIBUTED COMPUTING**

**Student Handout**

This module was initially developed by Carey, C.C., S. Aditya, K. Subratie, and R. Figueiredo. 1 May 2016. Project EDDIE: Modeling Climate Change Effects on Lakes Using Distributed Computing. Project EDDIE Module 4, Version 1. <http://cemast.illinoisstate.edu/data-for-students/modules/lake-modeling.shtml>. Module development was supported by NSF DEB 1245707 and ACI 1234983.

Learning Objectives:

* Set up and run the General Lake Model (GLM) in the R statistical environment to simulate lake thermal structure.
* Understand the structure and function of GLM configuration files, driver data, and output files.
* Modify the input meteorological data for one GLM model to simulate the effects of different climate scenarios on lake thermal structure.
* Interpret model output from GLM simulations to understand how changing climate will alter lake thermal characteristics.
* Use the GRAPLEr R package to set up hundreds of model simulations with varying input meteorological data, and run those simulations using distributed computing.
* Explore the application of distributed computing for modeling climate change effects on lakes.

Why this matters:

Lakes around the globe are experiencing the effects of climate change. Because it is difficult to predict how lakes will respond to the many different aspects of climate change (e.g., altered temperature, precipitation, wind, etc.), many researchers are using models to manipulate climate scenarios and simulate lake responses. Lake models provide a powerful tool for exploring the sensitivity of lake thermal structure characteristics to weather. In this module, you will learn how to set up a lake model and “force” the model with climate scenarios of your own design to examine how lakes may change in the future. While it is relatively easy to run one lake model on your own computer, it becomes more challenging to run hundreds of models because of the time-consuming nature of a high computational workload. To overcome this problem, we have developed an R package called GRAPLEr, which allows you to submit hundreds of model simulations through an interface in the R statistical environment, run those models efficiently and quickly using distributed computing tools, and then retrieve the model output. The GRAPLEr allows you to harness cyberinfrastructure tools commonly used in computer science to improve the speed of computing that are rarely used in ecology and freshwater sciences. Ultimately, using the GRAPLEr and similar tools will allow us to improve our understanding of climate change effects of lakes.

Outline:

1. Read the background and any assigned readings
2. Download R software onto your computer
3. Activity A: Plotting water temperatures from a lake model
4. Activity B: Develop a climate scenario, generate hypotheses, and model how the lake responds
5. Activity C: Using distributed computing to run hundreds of lake model simulations

Background:

How will lake thermal structure change in response to altered climate? To address this question, we are going to use an open-source hydrodynamic model called GLM (the General Lake Model) (Hipsey et al. 2014). The GLM model uses inputs of several different weather variables (including air temperature, solar radiation, wind, and precipitation) in a meteorological driver file as well as inflow and outflow data from streams entering and leaving the lake. GLM also has a configuration file, or ‘master script’ file called an nml file, which gives basic information about the lake (e.g., maximum depth, latitude, lake name, etc.) as well as instructions as to how the model should be run. These instructions include the simulation start and end times and dates, the time step of the model, and the names of the meteorological and inflow data files. From these inputs, the model will simulate water temperatures at many depths over time. For more information about GLM, see: http://aed.see.uwa.edu.au/research/models/GLM

Setting up R software and files:

You will need to first download R software onto your computer. You can use the file ‘EDDIE\_LakeModeling\_DownloadingRTutorial’ for help. Once you have R installed, all of the information you need for the modeling activities is embedded into the R scripts in the file ‘EDDIE\_LakeModeling\_R\_1May2016’ and chose the script file that’s appropriate for your computer (PC or Mac). Read through the annotation for each step so that you understand all of the lines of code that you are running.

**Activity A:** Plotting water temperatures from a lake model

1. With a partner, set up the GLM files and R packages on your computer (Objective 1 in the R script).
2. Access the files for an example lake and set up directories (a.k.a. ‘folders’) on your computer to run the model (Objective 2).
3. Run the model and explore the output (Objective 3).
4. Compare your model output with the observed field data for your lake (Objective 4).

**Activity B:** Develop a climate scenario, generate hypotheses, and model how the lake responds

1. With a partner, develop hypotheses as to how certain aspects of climate change may affect lake thermal structure.
2. Create a climate change scenario for your model lake to test the hypotheses, modify the input data accordingly, run the model, and analyze the output to determine how this scenario alters lake thermal structure (Objective 5).
3. After you have analyzed the model output, create some figures with your partner to present your model simulation and output to the rest of the class. Does the model output support or contradict your hypotheses?

**Activity C:** Using distributed computing to run hundreds of lake model simulations

1. Go step-by-step through the demonstration of the GRAPLEr package to examine how you can quickly set up many model simulations to have slightly different meteorological input files, submit these simulations, and return and analyze the output.
2. Once you have finished the demonstration, design your own simulation "experiment" with your partner and use the GRAPLEr to examine the offsets of a meteorological variable and magnitude of your choice.
3. Create some figures from your simulation and share them with the class.

References:

Hipsey, M.R., L.C. Bruce, and Hamilton, D.P. 2014. GLM - General Lake Model: Model overview and user information. AED Report #26, The University of Western Australia, Perth, Australia. 42pp.

Subratie, K., S. Aditya, R. Figueiredo, C.C. Carey, and P. Hanson. 2015. GRAPLEr: A distributed collaborative environment for lake ecosystem modeling that integrates overlay networks, high-throughput computing, and web services. PRAGMA Workshop on International Clouds for Data Science (PRAGMA-ICDS’15). arXiv e-prints 1509.08955, 8 p.