

Simulation Modeling of Lakes in Undergraduate and Graduate Classrooms Increases Comprehension of Climate Change Concepts and Experience with Computational Tools

Cayelan C. Carey¹ · Rebekka Darner Gougis²

Published online: 22 August 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Ecosystem modeling is a critically important tool for environmental scientists, yet is rarely taught in undergraduate and graduate classrooms. To address this gap, we developed a teaching module that exposes students to a suite of modeling skills and tools (including computer programming, numerical simulation modeling, and distributed computing) that students apply to study how lakes around the globe are experiencing the effects of climate change. In the module, students develop hypotheses about the effects of different climate scenarios on lakes and then test their hypotheses using hundreds of model simulations. We taught the module in a 4-hour workshop and found that participation in the module significantly increased both undergraduate and graduate students' understanding about climate change effects on lakes. Moreover, participation in the module also significantly increased students' perceived experience level in using different software, technologies, and modeling tools. By embedding modeling in an environmental science context, non-computer science students were able to successfully use and master technologies that they had previously never been exposed to. Overall, our findings suggest that modeling is a powerful tool for catalyzing student learning on the effects of climate change.

Keywords Simulation modeling · Climate change education · Hypothesis-testing

Introduction

Motivation for EDDIE Lake Modeling Module

Environmental scientists are increasingly analyzing large datasets of observations obtained through sensor networks and remote sensing, enabling new analyses and models of ecological phenomena (Hanson 2007; Porter et al. 2005; Weathers et al. 2013). Conducting these analyses and modeling as well as interpreting their results requires advanced skills in data manipulation, experimental design, quantitative reasoning, and data retrieval (Michener and Jones 2012; Langen et al. 2014; Schimel and Keller 2015). Despite the increasing importance of these skills, they are not commonly taught in undergraduate classrooms. To address this gap, Environmental Data-Driven Inquiry & Exploration (EDDIE) is a collaboration among ecologists and educators to develop stand-alone modular classroom activities for post-secondary classrooms using long-term (e.g., >10 years) and high-frequency (e.g., minute-scale) data (Carey et al. 2015). The goals of these modules, which follow the 5E learning cycle (Bybee et al. 2006), are to develop skills required to manipulate large datasets, conduct authentic investigations, develop reasoning about variability in data, engage students in scientific discourse as they explore large datasets, and foster sound ideas about the nature of environmental science research.

A primary aim of EDDIE is to teach students to use data analysis and modeling to investigate how climate change is altering ecosystems. Modeling is a critical tool for

Electronic supplementary material The online version of this article (doi:10.1007/s10956-016-9644-2) contains supplementary material, which is available to authorized users.

✉ Rebekka Darner Gougis
darner.gougis@ilstu.edu

¹ Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA

² School of Biological Sciences, Illinois State University, Normal, IL 61790, USA

environmental scientists because it allows them to study phenomena occurring at spatial and temporal scales for which we do not have observational data, as well as forecast the effects of future climate scenarios. Many ecosystem models are computationally intensive and written in scripting languages, so researchers need familiarity with different approaches of importing and exporting large datasets (e.g., using comma-separated values (CSV) or network common data form (NetCDF) files), different programming languages (e.g., C, Fortran, MATLAB, R), and different techniques for high-throughput computing. While it may be easy to run one model on a laptop computer, it becomes more challenging to run hundreds of model simulations because of the time-consuming nature of a high computational workload, requiring distributed computing approaches. These skills are generally not taught in most ecology undergraduate or graduate classrooms.

We developed an EDDIE module to examine how lakes around the globe are experiencing the effects of climate change, called “Modeling Climate Change Effects on Lakes Using Distributed Computing Module” (hereafter, Lake Modeling module; [module url]). Because it is difficult to predict how lakes will respond to the many different aspects of climate change (e.g., altered temperature, precipitation, wind), many aquatic ecologists are using lake models to manipulate weather 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, students learn how to set up a lake model and “force” the model with climate scenarios of their own design to examine how lakes may change in the future and interpret the output. To improve computational efficiency, students participating in the module also learn how to submit, retrieve, and analyze hundreds of model simulations through distributed computing technology embedded in an interface in the R statistical environment (Subratie et al. 2015).

As a result of this module, students are exposed to a suite of different computing tools and technologies. These include different file formats (Microsoft Excel spreadsheets, text files, and CSV files), a programming language (R statistical software; R Development Core Team 2015), an open-source hydrodynamic ecosystem model (GLM, the General Lake Model; Hipsey et al. 2014), and distributed computing techniques (overlay networks using peer-to-peer networking; Figueiredo et al. 2008; Ganguly et al. 2006; St. Juste et al. 2014). This module teaches students to harness cyberinfrastructure tools commonly used in computer science to improve the speed of computationally intensive Lake Modeling. The overarching goal of this module is to develop students’ computing and quantitative skills to

improve their understanding of how climate change is affecting lake ecosystems.

The Application of Modeling to the Learning Process

Modeling is integral to science, but engaging students in modeling also imparts cognitive benefits, allowing them to develop scientific knowledge that might not otherwise be realized. Modeling compels students to search for patterns in data, propose plausible hypotheses for causes of such patterns, and make evidence-based predictions (Stewart et al. 2005). When students have the opportunity to assess the utility of their models, reflection on earlier thinking is part of the basic process (Stewart et al. 2005). Students are naturally prompted to practice metacognitive skills, or the act of “monitoring, guiding and controlling one’s learning a problem-solving behavior” (Veenman 2011, p. 24), which profoundly supports conceptual development (Zohar and Dori 2011). Even young students who repeatedly engage in modeling practices can construct models that provide explanatory mechanisms, make predictions based on their models, and revise their models in light of new findings (Schwarz et al. 2009). As students gain more experience with modeling, their assessments of the quality of models shift from a binary perspective of “incorrect versus correct” toward a perspective in which they are able to evaluate a model’s ability to provide explanations for multiple aspects of the natural phenomenon being examined (Schwarz et al. 2009). In doing so, students come to understand how model building can help scientists make sense of natural processes and generate new scientific knowledge (Schwarz and White 2005; Schwarz et al. 2009).

The importance of modeling as both a scientific and a pedagogical tool has led the scientific education community to make it a central focus of science education, particularly through its incorporation into the Next Generation Science Standards (NGSS Lead States 2013). While the science education community embraces modeling as a scientific practice, the modeling typically employed in science classrooms is not the computer-based modeling that ecologists use to understand natural phenomena. Rather, there are domain-specific ideas of modeling among students; students tend to see the functional of models in biology as descriptive but predictive in the contexts of chemistry and physics (Krell et al. 2015). Students in K-12 grades do not conceptualize models in any domain as mathematical (Krell et al. 2015), which motivates us to examine additional skills that are needed to maneuver large sensor-based datasets and perform computer-based modeling that is at the heart of the “eco-informatics” revolution in ecology.

Climate Change as a Context for Learning

Climate change is an ideal context in which to engage students in modeling and exploration of large datasets gathered by sensor networks. Much of the climate change education literature focuses on students' perceptions of climate change and the plausibility of it being human induced. For example, Sinatra et al. (2012) examined relationships between motivational variables and college students' willingness to take action to mitigate climate change's effects. They identified openness to change and willingness to think deeply about issues as significant predictors of attitudinal change and an expressed willingness to change behavior. Others have investigated students' understanding of climate change mechanisms and consequences, often focusing on identifying misconceptions (Shepardson et al. 2013). Common misconceptions include a conflation of climate change with ozone depletion, pollution, and acid rain, in which causal mechanisms for the greenhouse effect and ozone depletion are often seen as the same phenomenon (Papadimitriou 2004). Shepardson et al. (2013) demonstrated how young students tend to perceive climate systems as unidirectional and linear, and they overemphasize the atmospheric component of climate systems. Students are often unable to distinguish climate and weather, leading them to cite short-term weather observations as evidence refuting long-term climatic changes (Lombardi and Sinatra 2012). Evidence suggests that this misconception can be improved with brief instruction that develops understanding of deep time, thereby improving students' perception of whether human-induced climate change is plausible (Lombardi and Sinatra 2012).

To our knowledge, only one study has looked at students' understanding of climate change in the context of lake ecosystems, and this study examined seventh graders' engagement in a climate change simulation with Lake Mead, which lead to improved student outcomes on pre-/post-simulation survey items related to understanding water conservation, the greenhouse effect, water flow, and weather versus climate (Nussbaum et al. 2015). Anderson (2012) provides a thorough review of the climate change education literature, concluding that comprehensive climate change education must address content knowledge of climate change in addition to environmental and social issues, disaster risk reduction, sustainable lifestyle decisions, and institutional policy. While we do not disagree with this broad approach, we argue that the climate change education literature is severely lacking in its exploration of how advanced content knowledge of climate change develops in novice scientists, particularly in how students develop understanding of climate change's effects on lake ecosystems. Here, our focus is on the effects of climate

change on lake physics—specifically, temperatures, thermal stratification, and mixing—because these physical variables provide the basis for all other climate change-induced effects on lake chemistry and biology.

Learning the Methods of Science

The science education community has diverged from defining science as “the scientific method” because it fails to capture the diversity of approaches to scientific inquiry and contributes to the development of misconceptions about what science actually is (McComas 1996). Students nevertheless have been taught and continue to believe the scientific method to be the hallmark of science (Miller et al. 2009). The Lake Modeling module provides an opportunity to engage students in more authentic practices of science, such as by demonstrating how data collection may precede hypothesis generation and testing and by using modeling to generate new scientific knowledge. Thus, we might expect the Lake Modeling module to change how students think about the methodologies of scientific investigations. Similarly, we might expect changes in students' ideas about how creativity is employed during science investigations, because the Lake Modeling module compels students to conceptualize plausible analyses with a preexisting dataset that would yield insights into effects of climate change, as well as generating their own data to force the model with different climate scenarios of their own creation.

It is noteworthy that EDDIE modules do not explicitly prompt students to reflect on the scientific inquiry they engage in during modules and students' preconceptions about the nature of scientific methodologies. This approach to teaching science methods is implicit, which assumes that by engaging in scientific inquiry, students will come to understand the methods of science (Khishfe and Abd-EL-Khalick 2002). Evidence from studies of younger students indicates that implicit approaches to teaching the methods of science are ineffective (Khishfe and Abd-El-Khalick 2002; Khishfe and Lederman 2006). The EDDIE Lake Modeling module is directed toward upper-level science majors and graduate science students, a population that is not often examined in the nature of science literature.

The current study seeks to document any improvements in the understanding of science methods that may occur as a result of the implicit approach used in EDDIE modules. We acknowledge greater learning gains may be plausible through an explicit-reflective approach involving direct instruction on science methods and guided reflection on their modeling experiences. However, because a major emphasis of the module is hypothesis-testing via modeling, rather than via classical experimentation as is usually emphasized in science courses, we consider it plausible that

this contrasting experience may elicit detectable changes in how students think about scientific methods.

Objectives

This study examined learning gains resulting from EDDIE's Lake Modeling module among upper-level biology majors and science graduate students. Three areas of learning were examined in pre-/post-surveys: (1) understanding of climate change effects on lakes, specifically focusing on the effects of altered climate on lake physics (i.e., temperature, mixing, and stratification), (2) students' perceptions of experience in using the technological tools and quantitative methods taught in the module (listed below), and (3) understanding of the nature of scientific methods as measured by the Student Understanding of Science and Scientific Inquiry (SUSI) scale (Liang et al. 2008). We used the surveys to answer the question: Were there changes in these three learning areas when we compare pre-module responses with post-module responses, and if so, how did changes vary between the undergraduate and graduate students?

Methods

EDDIE's Lake Modeling module was implemented in two workshops taught by the first author in September and October 2015. One workshop was taught to 10 undergraduate junior and senior biology majors interested in freshwater ecology on the campus of a large, doctorate-granting institution in the eastern USA. Another workshop was taught to 40 science graduate students interested in freshwater ecology at the 2015 Global Lakes Ecological Observatory Network (GLEON) annual conference in Chuncheon, South Korea. US students represented the majority of participants in both workshops, although a few students born outside of the USA participated in both workshops.

As much as possible, the instructor kept instruction consistent between the two workshops, which both lasted approximately half a day. The same graduate student assisted the lead instructor in helping students with computer modeling in both workshops, and all students used their own laptop computers in the workshops. Prior to both workshops, students received a handout that gave an overview of the module, instructions on how to download open-source R programming software (R Core Development Team 2015) onto their laptops, and R scripts for them to explore.

Module Overview

The Lake Modeling module consisted of five parts. First, the lead instructor presented an introductory PowerPoint

lecture on climate change effects on the thermal structure of lakes, an overview of the open-source lake hydrodynamics numerical simulation model used in the module (the General Lakes Model, GLM; Hipsey et al. 2014), and R software (see Supplements). Second, the students divided into pairs based on their laptop operating system (OS X or Microsoft Windows) to run the default version of the lake model in R and explore the output. Third, the student pairs developed a climate change scenario for their lake model and discussed hypotheses with their partners on how they expected their lake to respond *in silico*. The students then forced the lake with their climate scenario and analyzed the output to determine the effects of altered climate on lake thermal structure. Using figures the students created in R, each student pair gave short presentations on their model simulations to the rest of the workshop participants and discussed how their climate scenario affected their model lakes, specifically addressing if their original hypotheses were supported or disproven, and why. Fourth, after the instructor finished facilitating a discussion of the different scenarios and output, the instructor gave a demonstration of distributed computing software run in R to the students. Finally, the student pairs designed numerical simulation Lake Modeling experiments using GLM to run hundreds of model simulations with the same model parameterization and different climate driver data and analyzed the output. Throughout the module, the instructor encouraged the students to develop climate scenarios that were most interesting to them, instead of following a pre-defined list of possible scenarios the teacher developed.

In total, it took approximately 4 h for the students to complete the module activities, with additional time for breaks. Throughout the workshop, the instructor and graduate teaching assistant answered questions, debugged R code, and checked on the student pairs to ensure that everyone was engaged and able to complete the module activities. Most of the workshop time was allocated to the student pairs working together to run R scripts that contained code for setting a working directory, running GLM, modifying climate driver files, analyzing GLM output, and running the distributed computing software. We note that the instructors did not teach all of R programming in the module, but rather how to use R to run simulation models with heavily annotated code; the students that participated in this module did not write their own scripts but edited already-written ones the instructors prepared.

At the end of the PowerPoint lecture presentation and just before the students were allowed to begin working their partners, the instructor stated the six learning objectives of the module, which were referred to again at the end of the workshop. These learning objectives were (paraphrased): (1) set up and run GLM in the R programming environment to simulate lake thermal structure; (2)

understand the structure and function of GLM configuration files, climate driver data, and output files; (3) modify the input meteorological data for one GLM model to simulate the effects of different climate scenarios on lake thermal structure; (4) interpret model output from GLM model simulations to understand how changing climate will alter lake thermal characteristics; (5) use the distributed computing software in R to set up and run hundreds of model simulations with varying input meteorological data; and 6) explore the application of distributed computing for modeling climate change effects on lakes.

Data Collection

Volunteer participants were emailed a link to the pre-module survey 1 week prior to the beginning of both workshops. The first page of the online survey contained the informed consent document; informed consent was obtained from all individual participants included in the study. A similar survey was emailed to participants again within 1 week following the workshops.

The instruments used in this study measured three domains of knowledge: climate change effects on lakes, quantitative skills indicated by perceived experience in using a list of technological tools used in modeling, and an understanding of the nature of science. We measured understanding of climate change effects on lakes using four items (A–D). Climate Change Item A asked students to interpret a figure of lake temperatures over time and hypothesize what factors were responsible for changes in water temperature. Climate Change Item B asked students to predict how climate scenarios would affect lake heating. Climate Change Item C asked students to predict how climate scenarios would affect thermal stratification, and Climate Change Item D asked students to predict how lake responses to climate change would be context dependent on different lake characteristics. Each climate change item was multiple choice and contained a single correct response (see Electronic Supplementary Material). Scores across these four items were summed, and overall pre-/post-gains in understanding of climate change effects on lakes were compared using a split-plot analysis of variance (ANOVA), with time (pre/post) as a within-subjects factor and level (undergraduate vs. graduate) as a between-subjects factor. Due to the low sample size and internal reliability (Cronbach's $\alpha = 0.25$), a Cochran's Q test was also performed on each climate change item to examine differences in the probabilities of correct responses to each item before and after the workshops.

Understanding of the nature of science was measured using the Student Understanding of Science and Science Inquiry (SUSSI) instrument (Liang et al. 2008). The SUSSI

is comprised of six subscales that correspond to understanding of observation and inference, change of scientific theories, scientific laws versus theories, social and cultural influences on science, imagination and creativity in science, and methodology of scientific investigation. Students responded to each statement on a 1–5 agree/disagree Likert scale. Scores across SUSSI items were summed and overall pre-/post-gains in understanding of science and scientific inquiry were compared using a split-plot ANOVA, with time (pre/post) as a within-subjects factor and level (undergraduate vs. graduate) as a between-subjects factor. Due to the low sample size and reduced internal reliability (Cronbach's $\alpha = 0.64$) compared to the original reliability analysis (Liang et al. 2008), Wilcoxon signed-rank tests were also used to compare responses on individual SUSSI items before and after the workshop.

To measure students' perceived experience in various technological tools used in quantitative methods, students were asked, "On a scale from 1 to 5, with 1 indicating no prior experience whatsoever and 5 being very knowledgeable, how would you rank your experience level with _____?" followed by a list of quantitative tools, software, and technologies that included Excel software, CSV files (comma-separated values files), R software, computer programming, numerical simulation modeling, General Lake Model (GLM), distributed computing, and overlay networks. Students responded on a 1–5 Likert scale for each item.

We consider this measure a metric of their *perceived* experience level, because we know that they actually are more experienced with several of these technological tools after being taught the module, but this measure is to determine whether they feel more experienced with these tools after the module. This measure demonstrated high internal reliability (Cronbach's $\alpha = 0.89$). Scores across these items were summed and overall pre-/post-gains in perceived experience using these quantitative tools were compared using a split-plot ANOVA, with time (pre/post) as a within-subjects factor and level (undergraduate vs. graduate) as a between-subjects factor. Because we wanted to identify which technological tools drove any significant gains in perceived experience level, Wilcoxon signed-rank tests were also used to compare responses on each of these items before and after workshops.

For all survey instrument analyses, we interpreted statistical significance at $\alpha = 0.10$ to maximize our statistical power to detect whether any changes in pre- and post-module responses occurred (Quinn and Keough 2002). This α level was chosen because of our small sampling size ($n = 19$ respondents in total across both workshops) and the short amount of time students spent in the workshops. All statistical analyses of the survey data were conducted in R.

Results

Our results suggest that participation in the module significantly increased both undergraduate and graduate students' understanding about climate change effects on lakes (Fig. 1). In the analysis of variance, the interaction was nonsignificant, but when graduate and undergraduate groups were pooled across climate change items, pre–post-gains (pre-mean = 2.2 ± 0.55 , 95 % CI; post-mean = 3.0 ± 0.41 , 95 % CI) were statistically significant ($F_{1,17} = 6.55$, $p = 0.02$, partial eta squared (η^2) = 0.28, observed power = 0.67). For two of the four climate change items, the aggregated student population exhibited a significantly higher proportion of correct responses post-module than pre-module (Fig. 1a, b; Table 1). For the other two items, students either exhibited a nonsignificant increase in correct responses after the module (Fig. 1d) or no change (Fig. 1c), but no decreases in correct responses were observed.

Interestingly, the pattern of undergraduate and graduate responses differed; undergraduates exhibited a greater (significant) increase in correct responses after participating in the module for Climate Change Item A, whereas graduate students exhibited a greater increase (but not significant) in correct responses post-module for Items B and D. Student level did not appear to affect their ability to correctly answer the questions, as the undergraduates scored higher on some questions, whereas the graduates scored higher on others. In general, more than half of the students in both of the undergraduate and graduate classes correctly answered the questions in the post-module assessment.

Participation in the module also significantly increased students' perceived experience with different software, technologies, and modeling tools. In the analysis of variance, the interaction was significant, with both graduate (pre-mean = 21.4 ± 3.6 , 95 % CI; post-mean = 22.0 ± 4.0 , 95 % CI) and undergraduate (pre-mean = 12.6 ± 4.2 , 95 % CI; post-mean = 16.9 ± 4.7 , 95 % CI) students improving from pre- to posttest but undergraduates more so ($F_{1,17} = 4.0$, $p = 0.06$, partial $\eta^2 = 0.19$, observed power = 0.47). Wilcoxon signed-rank tests showed that the aggregated students' perceived experience level with CSV files, R software, the General Lake Model (GLM), distributed computing, and overlay networks was significantly higher post-module than pre-module (Table 2); these differences were largely driven by undergraduate responses, who reported a greater gain in experience with these tools than the graduate students.

Finally, participation in the module also affected student's conceptions of the nature of science, as measured by two individual items. Analysis of variance on the entire

instrument did not detect significant changes. However, statistically significant improvements were observed on two items (Table 3): (1) "Scientists' observations of the same event will be the same because observations are facts," and (2) "Scientific theories based on accurate experimentation will not be changed." Both undergraduates and graduates contributed to these gains, although neither group showed significant changes independently.

Discussion

Our findings and experience in the classroom both indicate that significant gains in improving students' understanding of climate change and perceived level of experience using modeling tools can occur with participation in this brief (~4 h) teaching module. Our results are encouraging because the modular and flexible format of the Lake Modeling teaching materials enables rapid dissemination and transfer of the module to other classrooms at both undergraduate and graduate levels. Overall, our findings suggest that modeling is a powerful tool for catalyzing student learning on the effects of climate change.

In both applications of the module, we observed significant improvement in undergraduate and graduate students' ability to correctly answer questions regarding climate change effects on lakes. Our results suggest that the undergraduate and graduate student populations responded to different aspects of the module because the two groups experienced different gains from the pre- to post-module on the four climate change questions. For example, while 30 % of the undergraduate students answered Climate Change Item A correctly on the pre-module assessment and 100 % correctly in the post-module assessment, the percent of graduate students correctly answering question A did not change before and after the module. The large improvement in the undergraduate students may be because they were less familiar with time series figures of water temperature before the module than the graduate students.

Correctly answering the climate change questions required high-order comprehension skills, interpreting data figures and scenarios, making predictions, and applying ecological understanding to new situations. The Lake Modeling module engages students in a variety of activities, such as exploring an example model, generating hypotheses, and evaluating concepts in light of modeling outcomes, explanation, justification, and interactive discussion. All of these activities provide students the opportunity to construct new knowledge that can facilitate sense-making in a transfer situation (Nokes-Malach and Mestre 2013), such as what students encountered on the post-module assessment. Interestingly, at the time of the

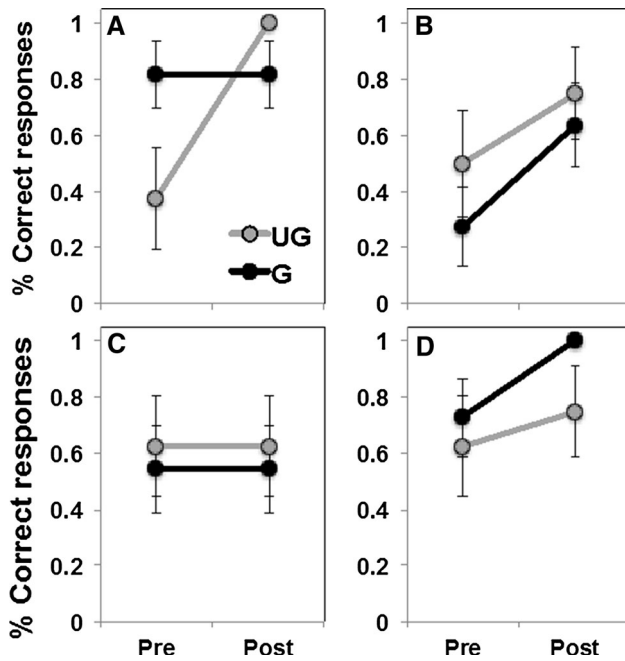


Fig. 1 Mean proportion (± 1 SE) of correct responses for undergraduate and graduate students on four climate change items. In Item A, students interpreted a figure and hypothesized factors responsible for changes in water temperature; in Item B, students predicted how climate scenarios would affect lake heating; in Item C, students predicted how climate scenarios would affect thermal stratification; and in Item D students predicted how lake responses to climate change would be context dependent on different lake characteristics

post-module assessment, undergraduate students showed higher proficiency than the graduate students on 3 out of 4 questions (Fig. 1), though we note that sample sizes for both populations were small and both student groups had varied exposure to these concepts prior to the module, likely influencing their responses. Regardless, it is

notable that there was no decrease in correctly answering the questions for either student group.

The students responded favorably to the *extension* phase of the module (Bybee et al. 2006), which challenged them to create a new climate scenario and develop hypotheses about how their climate scenario would affect their model lake in GLM. We noted that the students, regardless of education level, generally followed one of two patterns. They either explored how extreme conditions affected lake thermal structure, or they examined how future conditions predicted for a certain region would alter their model lake. For example, one group of undergraduate students accessed historical weather data from the US National Climatic Data Center (NCDC, <http://www.ncdc.noaa.gov>) and modified their climate driver data to compare the effects of severe hurricanes occurring in different seasons on lake thermal structure. Another group accessed downscaled climate predictions from the US National Aeronautics and Space Administration (NASA) NEX-DCP30 dataset (<http://climatedata.us>) and ran climate scenarios to simulate conditions for regions where they grew up for 2050 and 2100.

In both undergraduate and graduate classrooms, one of the most valuable components of the module was the discussion after each of the student teams presented the results from their climate scenarios to the class, which forced them to re-examine their earlier hypotheses of the scenario’s effects. This discussion required the students to articulate their justification for their initial hypotheses, evaluate whether the hypotheses were supported, and describe potential mechanisms for why that may or may not have occurred for their model lake. In both classroom experiences, this discussion engaged students in authentic scientific discourse, accomplishing one of the major goals of EDDIE.

Table 1 Statistical results of Cochran’s *Q* tests to examine differences in the probabilities of students correctly answering climate change items before and after the module

Climate change item	Population	Cochran’s <i>Q</i>	<i>p</i> value
A. Changes in water temperature	All students	3.57	0.057
	Undergraduates	5.00	0.025
	Graduates	0.00	1.00
B. Climate effects on lake heating	All students	4.50	0.03
	Undergraduates	2.00	0.16
	Graduates	2.67	0.10
C. Climate effects on stratification	All students	0.00	1.00
	Undergraduates	0.00	1.00
	Graduates	0.00	1.00
D. Lake characteristics	All students	1.80	0.18
	Undergraduates	0.33	0.56
	Graduates	2.00	0.16

Significant differences between pre- and post-module responses are highlighted in bold ($\alpha = 0.10$); undergraduate students ($n = 8$) and graduate students ($n = 11$) were assessed for Items A–C, although one graduate student skipped question D ($n = 10$)

Table 2 Statistical results from Wilcoxon signed-rank tests to measure differences in pre- and post-module responses of undergraduate students ($n = 8$) and graduate students ($n = 11$) to the question, “On a scale from 1 to 5, with 1 indicating no prior experience whatsoever and 5 being very knowledgeable, how would you rank your experience level with _____?”

Tool, software, or technology	Population	Test statistic	Two-tailed p value	Pre-module mean \pm 1 S.E.	Post-module mean \pm 1 S.E.
Excel software	All students	1.0	1.00	3.79 \pm 0.22	3.82 \pm 0.19
	Undergraduates	0.0	1.00	3.38 \pm 0.42	3.38 \pm 0.32
	Graduates	0.5	1.00	4.09 \pm 0.21	4.18 \pm 0.18
CSV files	All students	10.5	0.03	2.68 \pm 0.36	3.05 \pm 0.35
	Undergraduates	5.0	0.13	1.25 \pm 0.25	1.88 \pm 0.40
	Graduates	1.5	0.50	3.73 \pm 0.33	3.91 \pm 0.34
R software	All students	17.5	0.03	2.26 \pm 0.31	2.63 \pm 0.26
	Undergraduates	10.5	0.03	1.25 \pm 0.25	2.13 \pm 0.35
	Graduates	0.0	1.00	3.00 \pm 0.38	3.00 \pm 0.33
Computer programming	All students	3.5	0.69	2.16 \pm 0.26	2.26 \pm 0.21
	Undergraduates	3.0	0.25	1.38 \pm 0.38	2.64 \pm 0.37
	Graduates	-1.0	1.00	2.73 \pm 0.24	2.64 \pm 0.20
Numerical simulation modeling	All students	7.0	0.36	2.16 \pm 0.29	2.37 \pm 0.27
	Undergraduates	3.0	0.50	1.50 \pm 0.38	1.88 \pm 0.40
	Graduates	1.0	1.00	2.64 \pm 0.36	2.73 \pm 0.33
General Lake Model (GLM)	All students	27.5	0.002	1.68 \pm 0.27	2.37 \pm 0.27
	Undergraduates	10.5	0.03	1.25 \pm 0.25	2.25 \pm 0.45
	Graduates	5.0	0.13	2.00 \pm 0.40	2.45 \pm 0.34
Distributed computing	All students	10.5	0.03	1.42 \pm 0.18	1.74 \pm 0.18
	Undergraduates	3.0	0.25	1.45 \pm 0.38	1.73 \pm 0.37
	Graduates	3.0	0.25	1.38 \pm 0.16	1.75 \pm 0.19
Overlay networks	All students	14.0	0.06	1.21 \pm 0.12	1.36 \pm 0.19
	Undergraduates	5.0	0.13	1.25 \pm 0.25	1.36 \pm 0.40
	Graduates	2.5	0.63	1.18 \pm 0.12	1.36 \pm 0.15

Significant differences between pre- and post-module responses are highlighted in bold ($\alpha = 0.10$)

Table 3 Statistical results from Wilcoxon signed-rank tests to assess differences in pre- and post-module responses of undergraduate students ($n = 8$) and graduate students ($n = 11$) to items from the Student Understanding of Science & Scientific Inquiry (SUSSI) instrument

SUSSI Item	Population	Test statistic	Two-tailed p value	Pre-assessment mean \pm 1 S.E.	Post-assessment mean \pm 1 S.E.
Scientists' observations of the same event will be the same because observations are facts	All students	14.0	0.06	3.68 \pm 0.19	4.06 \pm 0.17
	Undergraduates	4.5	0.38	3.63 \pm 0.18	4.00 \pm 0.27
	Graduates	3.0	0.25	3.73 \pm 0.30	4.10 \pm 0.23
Scientific theories based on accurate experimentation will not be changed	All students	18.0	0.035	3.42 \pm 0.19	3.83 \pm 0.12
	Undergraduates	3.0	0.26	3.63 \pm 0.18	4.00 \pm 0.00
	Graduates	7.5	0.18	3.27 \pm 0.30	3.70 \pm 0.21

Significant differences between pre- and post-module responses are highlighted in bold; items with nonsignificant differences ($p \geq 0.10$) are not shown

In addition to improving their understanding of climate effects on lake ecosystems, the module also resulted in significant learning gains in students' perceived experience

level with modeling and several different computational and analytical tools. Participation in the module significantly increased undergraduates and graduate students'

perceived experience level with CSV files, R software, the GLM model, distributed computing, and overlay networks. For most of these tools, the undergraduate students exhibited greater gains than the graduate students, likely because they had less exposure to these tools prior to participating in the module, which was reflected in the pretest mean scores. Given the high internal reliability of these items, it is plausible that some items could be measuring the same construct. Upon examination of the inter-item correlation matrix, we found that measures of students' perceived experience level with CSV files and R software were highly correlated (>0.7), as were the items for distributed computing and overlay networks. Thus, it is plausible that these correlating items are measuring the same construct from the students' point of view, such as if a student is conflating overlay networks with distributed computing. Nonetheless, we think that it is noteworthy that a 4-hour module was able to catalyze significant increases in students' perceived experience level, especially for computational tools that are rarely taught in the classroom at either education level but germane to the field of ecology.

We posit that workshops or intensive short-term activities, such as the Lake Modeling module, that use computational tools framed around environmental science issues are effective in improving quantitative and computational literacy of non-computer science students because they generate intellectual need for the computing. In the context of such a learning task, a student's intellectual need results from his or her realization that the ability to solve a compelling problem, such as predicting climate change effects on lakes, requires the development and utilization of efficient methodologies, such as computational modeling (Fuller et al. 2011). In order to generate intellectual need, the problem posed to students in the task must be intrinsic to the student (Lim 2009). Since most students in our study are preparing to be scientists and are thus inherently interested in climate change and its effects on ecosystems, the ability to predict how climate scenarios affect lakes is an intrinsic problem to these students. Therefore, it sufficiently compels the development and use of computational modeling. When students are introduced to new methods and tools out of context, instructors often fail to help them realize the utility of those methods and tools in solving problems that are intrinsic to the students, so they feel intellectually aimless (Fuller et al. 2011). The Lake Modeling module, however, began with an objective, understanding the effects of climate change, which students already held as a personal goal. Thus, they understood the premise and the purpose of the activity, resulting in their attentive engagement throughout the entire module.

We note that students exhibited the greatest gains in their experience level in GLM, a numerical simulation

model. Numerical simulation is a powerful computational tool for testing the effects of different complex scenarios, but is rarely taught in most environmental science classrooms at either the undergraduate or graduate level. For students studying the effects of climate change on ecosystems, numerical simulation provides a way for students to both analyze and visualize how different aspects of climate change (e.g., altered air temperature, wind, precipitation, humidity; IPCC 2014) may interact over time series, which would be impossible to do analytically. Moreover, numerical simulation modeling provides a way for students to generate hypotheses, create driver data to test their hypotheses, and easily run a model to test whether their hypothesis is supported. Our findings suggest that numerical simulation may be a useful tool for engaging students in exploration of climate change, which in turn increased their experience level in GLM. Again, we note that there was no significant decrease in students' perceived experience level in any of these computational tools from the pre- to posttest.

With emphasis on the utility of modeling and effects of climate change on lakes, there was little instructional attention paid to how the methods utilized in this module differ from the canonized scientific method. Nonetheless, students, particularly undergraduates, experienced gains on items related to scientific observations and the tentativeness of scientific knowledge. While speculative, it is plausible that engaging in computational modeling implicitly emphasized the nature of observations and the inferences that can be drawn from them and how modeling can produce new scientific knowledge. These marginal gains further support the notion that the implicit model of teaching nature of science concepts is not likely to be effective in teaching the breadth of nature of science concepts (Khishfe and Abd-EL-Khalick 2002). However, EDDIE modules, which engage students in science methods that do not fit cleanly into the scientific method, provide opportunities to mix integrated and nonintegrated approaches to the teaching of the nature of science and scientific methods, which is likely to be more effective in teaching the nature of science (Khishfe and Lederman 2006). Specific to the Lake Modeling module, one phase of the lesson compels student pairs to design numerical simulation Lake Modeling experiments and encourages them to develop climate scenarios they find most interesting, not adhering to scenarios they think the instructor is anticipating or would have developed himself/herself. Conceptualization of such climate scenarios requires much creativity and imagination, because it involves developing climate scenarios that have not been directly experienced before and extrapolating what effects those scenarios might have on lakes, given what the students have learned about

how lakes behave in the climate conditions explored thus far. Further, the entire module is built around the notion of hypothesis-testing using modeling, which differs from the hypothesis-testing they are most often confronted with in typical science coursework (e.g., hypothesis-testing via classical experimentation). The Lake Modeling module provides students with an exemplary experience demonstrating how scientific methods do not adhere to the canonized single scientific method and how imagination and creativity are necessary to conduct productive science. The module provides instructors opportunity to pose reflective questions to students on their experience conducting scientific inquiry within the module and confront common misconceptions about what counts as legitimate scientific practice.

The goal of EDDIE is to create teaching modules that are modular in structure and can be easily applied to classrooms at different student levels. Data from this study, as well as assessment data from three other EDDIE modules taught in different classrooms in the USA that spanned introductory biology majors to graduate students in freshwater ecology (Carey et al. 2015), demonstrate that this pedagogical approach can be successful. While our data indicate that the undergraduate and graduate students responded to different aspects of the module, both student groups showed significant increases in their understanding of climate change effects on lakes, as well as increases in their perceived experience in using modeling tools.

Our data also suggest that engaging students in hands-on modeling may substantially increase their appreciation for modeling as a science methodology. The three earlier EDDIE modules focused on exploration and analysis of long-term lake ice-off data and high-frequency lake-mixing and metabolism data (Carey et al. 2015). Here, we challenged students to move beyond analyses of data they were given a priori from an instructor to create hypotheses, use a fairly sophisticated model to test their hypotheses, and then analyze output data they generated themselves to see whether their hypotheses were supported. While computational modeling is used in graduate dissertations and advanced research projects, it is rarely applied in undergraduate science classrooms, thereby preventing undergraduates from realizing the potential that modeling could hold for pursuing graduate work and future careers.

Limitations and Future Research

As noted above, our interpretation of the assessment data is limited by our small sample size across two workshops. We strongly recommend that the Lake Modeling module be taught in multiple classrooms to assess the generalizability of our findings, as well as determine how initial student

experience level may mediate their gains in perceived experience level and climate change understanding. Moreover, it would be useful to examine how students respond to modeling other effects of climate change on lakes, such as chemistry and biology, and whether they yield similar results to models focused on lake physics. All of the teaching materials used in the module are available online at [module url], and we encourage other instructors to adapt them to their classrooms for building both quantitative skills and climate change understanding for their students.

Conclusions

Here, we found that use of computational modeling to explore the effects of climate change may be a new instructional strategy that can stimulate learning and improved comprehension of complex topics that are difficult to understand from static datasets. Furthermore, exposing undergraduates to modeling prior to graduate school lessens any potential intimidation that might prevent them from using modeling methods in their careers. By embedding modeling and other computational tools (e.g., distributed computing, overlay networks) in an environmental science context, non-computer science students were able to successfully use and master technologies that may otherwise be overwhelming. Consequently, using modular teaching materials such as the Lake Modeling module in the classroom may equip students with new tools for better understanding and predicting how complex environmental challenges, such as climate change, are altering the ecosystems they live in.

Acknowledgments We thank the amazing undergraduate students at Virginia Tech and graduate students in GLEON who participated in the Lake Modeling module and provided data for this study. We are grateful to the entire Project EDDIE team, especially Catherine O'Reilly, for their support and assistance. Saumitra Aditya, Ken Subratie, Renato Figueiredo, and Paul Hanson developed the distributed computing overlay network tools for the module as part of the GRAPLE (GLEON Research and Lake PRAGMA Expedition) team, and Jon Doubek and Kate Hamre provided invaluable assistance teaching the module. We appreciate administrative support provided by CeMaST (Center for Mathematics, Science, and Technology) at Illinois State University. This work was financially supported by grants from the National Science Foundation (DEB 1245707 and ACI 1234983).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anderson A (2012) Climate change education for mitigation and adaptation. *J Educ Sustain Dev* 6(2):191–206
- Bybee RW, Taylor JA, Gardner A, Van Scotter P, Powell JC, Westbrook A, Landes N (2006) The BSCS 5E instructional model: origins, effectiveness, and applications. BSCS, Colorado Springs
- Carey CC, Gougis RD, Klug JL, O'Reilly CM, Richardson DC (2015) A model for using environmental data-driven inquiry and exploration to teach limnology to undergraduates. *Limnol Oceanogr Bull* 24(2):32–35
- Figueiredo RJ, Boykin PO, Juste PS, Wolinsky D (2008) Integrating overlay and social networks for seamless p2p networking. In: IEEE 17th workshop on enabling technologies: infrastructure for collaborative enterprises, 93–98
- Fuller E, Rabin JM, Harel G (2011) Intellectual need and problem-free activity in the mathematics classroom. *Int J Stud Math Educ* 4(1):80–114
- Ganguly A, Agrawal A, Boykin PO, Figueiredo R (2006) IP over P2P: enabling self-configuring virtual IP networks for grid computing. In: Proceedings of the 20th international parallel and distributed processing symposium. Rhodes Island, Greece
- Hanson PC (2007) A grassroots approach to sensor and science networks. *Front Ecol Environ* 5(7):343
- Hipsey MR, Bruce LC, Hamilton DP (2014) GLM—General Lake Model: Model overview and user information. AED Report #26. The University of Western Australia, Perth, Australia. <http://aed.see.uwa.edu.au/research/models/AED/>
- IPCC (2014) Climate change 2014: impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds) Cambridge University Press, New York
- Khishfe R, Abd-EL-Khalick F (2002) Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of the nature of science. *Calif J Sci Educ* 9(1):119–173
- Khishfe R, Lederman N (2006) Teaching nature of science within a controversial topic: integrated versus nonintegrated. *J Res Sci Teach* 43(4):395–418
- Krell M, Reinisch B, Kruger D (2015) Analyzing students' understanding of models and modeling referring to the disciplines biology, chemistry, and physics. *Res Sci Educ* 45(3):367–393
- Langen TA, Mourad T, Brant BW, Gram WK, Abraham BJ, Fernandez DS, Carroll M, Nuding A, Balch JK, Rodriguez J, Hampton SE (2014) Using large public datasets in the undergraduate ecology classroom. *Front Ecol Environ* 12(6):362–363
- Liang LL, Chen S, Chen X, Kaya ON, Adams AD, Macklin M, Ebenezer J (2008) Assessing preservice elementary teachers' views on the nature of scientific knowledge: a dual-response instrument. *Asia-Pac Forum Sci Learn Teach* 9(1):1–20
- Lim KH (2009) Provoking intellectual need. *Math Teach Middle Sch* 15(2):93–99
- Lombardi D, Sinatra GM (2012) College students' perceptions about the plausibility of human-induced climate change. *Res Sci Educ* 42(2):201–217
- McComas WF (1996) Ten myths of science: reexamining what we think we know. *Sch Sci Math* 96(1):10
- Michener WK, Jones MB (2012) Ecoinformatics: Supporting ecology as a data-intensive sciences. *Trends Ecol Evol* 27(2):85–93
- Miller MCD, Montplaisir LM, Offerdahl EG, Cheng F, Ketterling GL (2009) comparison of views of the nature of science between natural science and nonscience majors. *CBE-Life Sci* 9(1):45–54
- NGSS Lead States (2013) Next generation science standards: for states, by states. The National Academies Press, Washington, DC
- Nokes-Malach TJ, Mestre JP (2013) Toward of model of transfer as sense-making. *Educ Psychol* 48(3):184–207
- Nussbaum EM, Owens MC, Sinatra GM, Rehmat AP, Cordova JR, Ahmad S, Harris FC Jr, Dascalu SM (2015) Losing the lake: simulations to promote gains in student knowledge and interest about climate change. *Int J Environ Sci Educ* 10(6):789–811
- Papadimitriou V (2004) Prospective primary teachers' understanding of climate change, greenhouse effect, and ozone layer depletion. *J Sci Educ Technol* 13(2):299–307
- Porter J, Arzberger P, Braun HW, Bryant P, Gage S, Hansen T, Hanson P, Lin C-C, Lin F-P, Kratz T, Michener W, Shapiro S, Williams T (2005) Wireless sensor networks for ecology. *Bioscience* 55(7):561–572
- Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, New York
- Schimel D, Keller M (2015) Big questions, big science: Meeting the challenges of global ecology. *Oecologia* 177(4):925–934
- Schwarz CV, White BY (2005) Metamodeling knowledge: developing students' understanding of scientific modeling. *Cogn Instr* 23(2):165–205
- Schwarz CV, Reiser BJ, Davis EA, Kenyon L, Acher A, Fortus D, Shwartz Y, Hug B, Krajcik J (2009) Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *J Res Sci Teach* 46(6):632–654
- Shepardson DP, Roychoudhury A, Hirsch A, Niyogi D, Top SM (2013) When the atmosphere warms it rains and ice melts: seventh grade students' conceptions of a climate system. *Environ Educ Res* 20(3):333–353
- Sinatra GM, Kardash CM, Taasoobshirazi G, Lombardi D (2012) Promoting attitude change and expressed willingness to take action toward climate change in college students. *Instr Sci* 40(1):1–17
- St. Juste P, Jeong K, Eom H, Baker C, Figueiredo R (2014) TinCan: user-defined P2P virtual network overlays for ad-hoc collaboration. *ICST Trans Collab Comput* 14(2):e4
- Stewart J, Cartier J, Passmore C (2005) Developing understanding through model-based inquiry. In: Donovan MS, Bransford JD (eds) How students learn. National Research Council, Washington, DC, pp 515–565
- Subratie K, Aditya S, Figueiredo R, Carey CC, Hanson P (2015) GRAPLER: a distributed collaborative environment for lake ecosystem modeling that integrates overlay networks, high-throughput computing, and web services. arXiv e-prints 1509.08955, 8 p. <http://adsabs.harvard.edu/abs/2015arXiv150908955>
- R Core Development Team (2015) R: a language and environment for statistical computing (version 3.2.2). R Foundation for Statistical Computing, Vienna, Austria. [Software]. <https://www.R-project.org/>
- Veenman MVJ (2011) Metacognition in science education: Definitions, constituents, and their intricate relation with cognition. In: Zohar A, Dori YJ (eds) Metacognition in science education: trends in current research. Springer, New York, pp 21–36
- Weathers K, Hanson PC, Arzberger P, Brentrup JA, Brookes J, Carey CC, Gaiser E, Hamilton DP, Hong GS, Ibelings BW, Istvanovics V, Jennings E, Kim B, Kratz TK, Lin F-P, Muraoka K, O'Reilly CM, Piccolo C, Ryder E, Zhu G (2013) The Global Lake Ecological Observatory Network (GLEON): the evolution of grassroots network science. *Limnol Oceanogr Bull* 22:71–73
- Zohar A, Dori YJ (2011) Metacognition in science education: trends in current research. Springer, New York