Effectively Extending Computational Training Using Informal Means at Larger Institutions

Dhruva K. Chakravorty
High Performance Research Computing
Texas A&M University
Texas, USA
chakravorty@tamu.edu

Marinus "Maikel" Pennings
High Performance Research Computing
Texas A&M University
Texas, USA
pennings@tamu.edu

Honggao Liu
High Performance Research Computing
Texas A&M University
Texas, USA
honggao@tamu.edu

Zengyu "Sheldon" Wei
High Performance Research Computing
Texas A&M University
Texas, USA
lele0027@tamu.edu

Dylan M. Rodriguez
High Performance Research Computing
Texas A&M University
Texas, USA
dylan@tamu.edu

Levi T. Jordan
High Performance Research Computing
Texas A&M University
Texas, USA
ljordan56@tamu.edu

Donald "Rick" McMullen
High Performance Research Computing
Texas A&M University
Texas, USA
mcmullen@tamu.edu

Noushin Ghaffari
High Performance Research Computing
Texas A&M University
Texas, USA
nghaffari@tamu.edu

Shaina D. Le
High Performance Research Computing
Texas A&M University
Texas, USA
sle1019@tamu.edu

ABSTRACT

Short courses offered by High Performance Computing (HPC) centers offer an avenue for aspiring Cyberinfrastructure (CI) professionals to learn much-needed skills in research computing. Such courses are a staple at universities and HPC sites around the country. These short courses offer an informal curricular model of short, intensive, and applied micro-courses that address generalizable competencies in computing as opposed to content expertise. The degree of knowledge sophistication is taught at the level of below a minor and the burden of application to domain content is on the learner. Since the Spring 2017 semester, Texas A&M University High Performance Research Computing (TAMU HPRC) has introduced a series of interventions in its short courses program that has led to a 300% growth in participation. Here, we present the strategies and best practices employed by TAMU HPRC in teaching short course modules. We present a longitudinal report that assesses the success of these strategies since the Spring semester of 2017. This data suggests that changes to student learning and a reimagination of the tiered instruction model widely adopted at institutions could be beneficial to student outcomes.

CCS CONCEPTS

• CS→Computer Science; • Cybertraining→training on using cyberinfrastructure; • HPC→high performance computing

Keywords

HPC training, broadening participation, assessment strategies, best practices, diversity, computational thinking, tiered instruction

DOI: https://doi.org/10.22369/issn.2153-4136/10/1/7

1. INTRODUCTION

Research efforts in STEAM (Science, Technology, Engineering, Art, and Mathematics) have significantly benefited from the rapid growth of computational capacity and the extensive use of data-analytics tools. The rapid proliferation of these methods has brought about an urgent need to train researchers who can effectively incorporate field-relevant computational tools and methods in their research workflows. In fact, developing computational and programming competency in the future science, technology, engineering, and mathematics workforce is a core component of the National Strategic Computing Initiative and the

National Science Foundation vision for Cyberinfrastructure for the 21st century [NSF NSCI]. In stark contrast to a global needs for a computationally-trained workforce, a vast majority of graduate students have limited exposure to computing. Indeed, during the NSCI presentations at SuperComputing 17 (SC17, Denver, CO) Irene Qualters (Director of the Office of Advanced Cyberinfrastructure, National Science Foundation NSCI Driver) discussed the need to train students to write coherent code [NSF HER]. There is a rapidly growing need to identify strategies to successfully introduce this population of students to computational methods and approaches [Lu 2009, NSF-Research, TRC 2014, Wing 2008, Yadav 2011].

Short courses and tutorials provided by HPC units remain stalwarts of informal education at the collegiate level. HPC-led short courses provide researchers with much needed technical information required for research and fill a void for student, staff, and faculty professional development that is not provided in a formal educational setting. Such courses further the institution's academic mission while simultaneously addressing the research computing needs of users who rely on these facilities. Unlike traditional credit-bearing courses that need to be approved at the department, college, and university level, an HPC unit can launch a short course in as little as two weeks. Furthermore, HPC units are constrained only by the expertise of their staff. While the importance of in-person training exercises cannot be stressed enough, "live" online training events organized at the regional or national levels are also effective. Events such as the XSEDE Big Data workshops and the Peta Scale Institute allow HPC units across the country to provide training on specialized topics that may go beyond local expertise at any specific site.

HPC-led courses are dynamic in nature. Owing to variations in the availability of expertise and researcher needs, HPC units have adopted different models of user training. At a rudimentary level, HPC short course offerings traditionally include courses that provide an introduction to operating systems on the cluster (Linux), cluster utilization (schedulers and file structures), and interpreted languages (Perl or Python). Larger HPC centers offer courses that included parallel programming paradigms (Open MP and MPI), rudimentary bioinformatics job-submission interfaces (Galaxy), and perhaps software applications as well (Abaqus or AMBER). As many branches of science have adopted large-scale computing, the HPC user profile has changed in recent years. We now offer additional courses that cover the use of data analysis toolkits (MATLAB, SAS and R), interfacing interpreted languages with data analysis packages (MySQL for Python users), and machine learning frameworks (TensorFlow and Caffe). This change in course offerings has been complemented by the gradual adoption of HPC course materials in into the formal classroom space. For example, TAMU HPRC does not offer the standard course that introduced Galaxy to our users. This material is now covered in the BIOL 647, "Digital Biology", a credit-bearing course taught by Prof. Rudolfo Aramayo with assistance from TAMU HPRC. Conversely, we have witnessed a significant growth in interest in our Python offerings because the traditional Computer Science course on Python is no longer approved for a graduate student's degree plan.

Researcher participation in HPC-led courses can be remarkably different at various institutions. This is surprising, considering that the topics covered in HPC courses and the computing needs of researchers largely remain the same across institutions of a similar size. Furthermore, the best practices in informal education are also well documented. The typical factors attributed to such variations are the instruction models used in teaching HPC courses, the amount of technology used in training, the length of the courses, the frequency of course offerings, the location of course offerings, the composition of research projects at the university, advertising information about the courses to the research community, student preparation at the undergraduate level, formal courses offered at the institution, and involving faculty in HPC instruction. While a number of interventions are possible to address these factors, there remains a dearth of quantitative data about the effects of these interventions on researcher participation in literature.

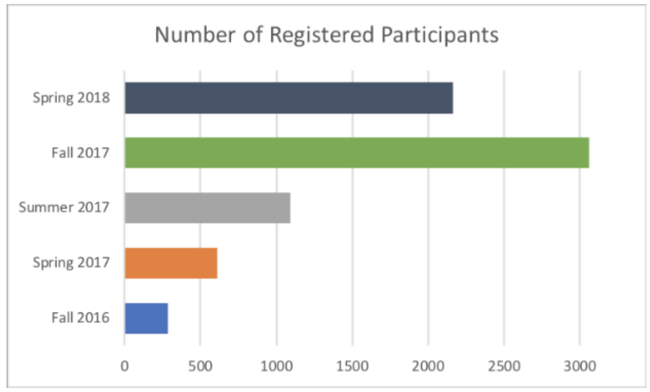

Figure 1. Number of registered participants in TAMU HPRC short courses from Fall 2016 through Spring 2018. The number of registered participants is not adjusted for hours of instruction, or the number of courses offered in each semester.

In this paper, we present a report on the effects of introducing curricular interventions on researcher participation in TAMU HPRC short courses program since Fall 2016. In the subsequent sections of this paper, we present quantitative data from our short course program along with details of our current offerings. We next describe a list of interventions that were introduced to our short courses program over the last four semesters. The paper next describes our efforts toward assessing the success of these courses on student learning using evaluations. We finally discuss the lessons learned over the previous year and summarize our findings in conclusion.

2. TAMU HPRC SHORT COURSES

TAMU HPRC short courses use active learning techniques and rely on HPRC staff expertise for content development. The courses are traditionally structured on a tiered instruction model (TIM) [Adams 2003, Tomlinson 1999, OME 2005, and OME 2013]. The tiered instruction approach provides lessons at different ability levels or areas of interest for a diverse learning community. Students and researchers using HPC resources come from varied academic and research backgrounds. In addition, these researchers may have different levels of exposure to computing and will require diverse computing skill sets to meet their research needs. In a typical TIM approach, the vast majority of learners would first participate in foundational courses, with a gradual drop-off as medium to higher level topics are approached. The TIM, however, faces challenges from advances in technology that eliminate the need for certain foundational courses and popular advanced offerings, such as CUDA or Machine Learning, that appeal to a wide range of researchers. These short courses are offered free of charge in an in-person format at the TAMU main campus in College Station, TX. These courses are also offered "live" via WebEx to an online audience that includes participants from a number of universities in the United States (including Puerto Rico). Figure 1 shows the number of registered participants in TAMU HPRC courses since Fall 2016. For the purposes of this paper, we use participation data from Fall 2016 to describe a baseline that is compared to subsequent semesters where curricular and technical interventions were implemented.

3. IMPLEMENTED INTERVENTIONS

A number of technical, curricular, and engagement interventions have been systematically applied on a semester-wise basis. These include greater visibility for the courses, engaging students with active learning methods, better advertising, and retaining student interest via our HPRC seminar series. These steps have been simultaneously complemented by improved documentation on our website and wiki. Table 1 describes a series of interventions that were implemented in the TAMU HPRC short course offerings starting in Spring 2017. Taking a semester-by-semester approach as opposed to implementing all interventions in one semester, allowed us to quantify the effects of each semester. This approach also lets one to refine each intervention individually while allowing a short-staffed operation to adjust to the changes in schedules. While these interventions have overall contributed to making our courses more accessible, we have seen that our evolving online platform has had the largest effect on participation in our courses.

Table 1. Interventions introduced to TAMU HPRC short courses since Fall 2016. Interventions were carried into following semesters unless otherwise noted.

Semester	Interventions
Fall 2016	2-hour long lecture format courses. Hands-on exercises were not included.
	Certificates of attendance provided to attendees.
Spring 2017	2-hour long lecture format courses.
	Printed flyers distributed across campus.
	All courses slides were made available online in a standard format post-course.
	Handouts offered to students.
	Surveys collected via email.
	Seminal course on databases offered.
Summer 2017	WebEx introduced and online registration systems tested.
	Multiple courses offered on the same day
	Courses co-located and advertised with research computing event and open to REU students visiting TAMU
Fall 2017	Registrations standardized via Google forms interface.
	Offered seminal course on data management practices new to HPC training nationwide.
	Courses are first advertised to TAMU HPRC users and then to the entire TAMU community via campus email
	Introduced a new course, titled “Introduction to R”. Python offerings increased to include w courses.
	Handouts and PowerPoint presentations offered pre-class online.
Fall 2017 continued	To avoid issues with user registrations on HPRC systems, virtual machines were used for short course support.
	All courses were broadcast via WebEx.
	Certificates were restricted to in-person attendees alone.
	Courses offered three days a week at two different locations near Engineering departments and biology/life sciences departments. These two locations were selected to ensure convenient commute for the participants.
	Partnered with the Laboratory for Molecular Simulation to offer new courses.
	Typical course length was 1.5 hours.
	Interactive exercises introduced.
Spring 2018	Surveys collected in-person on paper on conclusion of the courses.
	Classes offered on an all-day Friday setting at a single location near engineering and science departments.
	All courses offered in 3-hour format with a 10-minute break.
	All courses use active learning methods.
	Courses were recorded for future ADA compliant online courses.
	Introduction to Galaxy HPRC course discontinued. Supported BIOL647 “Digital Biology” a credit-bearing course.
	Offered training support to formal courses at TAMU.
	Standardized format to support XSEDE online workshops/courses.
	Open-on-demand shell access used in lieu of Moba-X-term and Putty during training.
	Offsite in-person training offered at other universities.
	Reports and analytics on short courses were prepared.
	Employed analytics to make decisions
	Machine Learning/Artificial course bouquet was offered to complement AI/ML support push by HPRC.

4. GROWTH IN PARTICIPATION

As described above Texas A&M HPRC offered a number of courses in Spring 2018. These courses and workshops were offered on a Friday morning and afternoon schedule to maximize the opportunities for researchers to attend these courses. A complete listing of short courses offered in Spring 2018 along with the number of registered participants in each course are provided in Table 2.

Table 2. List of TAMU HPRC short courses offered in Spring 2018. The courses are listed in the order in which they were

offered. The number of registered participants includes both online and in-person attendees.

Course Name	Registered Participants
Introduction to Linux	118
Data Management Practices	76
Workshop - Introduction to Linux	65
Introduction to HPRC Clusters	98
TAMU Open on Demand Portal	17
Deep Learning with TensorFlow	226
Molecular Modeling Workshop	30
Introduction to Python	217
Introduction to Scientific Python	190
Introduction to MATLAB	103
Python for MATLAB Users	89
Introduction to Perl	89
Introduction to Databases	108
Modern Computational Physics	11
Introduction to CUDA	43
Introduction to MATLAB Parallel Toolbox	10
Software Carpentry - Git, Shell & R	40
Using LS Dyna	18
Code Parallelization Using OpenMP	35
Code Parallelization Using MPI	30
Introduction to NGS	47
Introduction to NGS Assembly	45
Linux and Cluster Usage (TAMU Galveston)	18
Introduction to NGS Metagenomics	22
Introduction to NGS RADSeq/GBS	23
Introduction to the R programming language	136
Introduction to Fortran	57
Machine Learning and Deep Learning with MATLAB	183
XSEDE Big Data Workshop	40

4.1 Impact of the Academic Year

While we observed significant growth in participant registrations since Fall 2016, we recorded the highest number of participants in Fall 2017, when three HPRC courses were offered each week on a Tuesday, Wednesday, Friday schedule. Each course was taught for 90 minutes. This matches traditional expectations of the academic year, as new graduate students traditionally enroll at the university in the Fall semester. In Spring 2018, a number of these courses were consolidated to offer two courses each week that were taught for three hours each on a Friday morning and afternoon schedule. It is important to note that while the number of offered courses was reduced, the increase in instruction hours more than compensated for this effect. Furthermore, in Spring 2018, TAMU HPRC taught portions of formal graduate level courses that relied on the use of HPRC resources. These co-taught training models helped us strengthen ties with faculty and freed up time on our training program allowing us to offer new courses, a Software Carpentry series and support XSEDE workshops. Consequently, the total number of participants when adjusted for each hour of instruction represented a slight increase in Spring 2018 as compared to Fall 2017. This is surprising, as there were fewer new students to the Texas A&M campus in Spring 2018 as compared to the Fall 2017 semester. An additional compensating factor could be that existing graduate students at TAMU who had not previously enrolled in the HPRC short course program registered for the offerings in Spring 2018. We anticipate that participation data from Fall 2018 will bring clarity to this discussion.

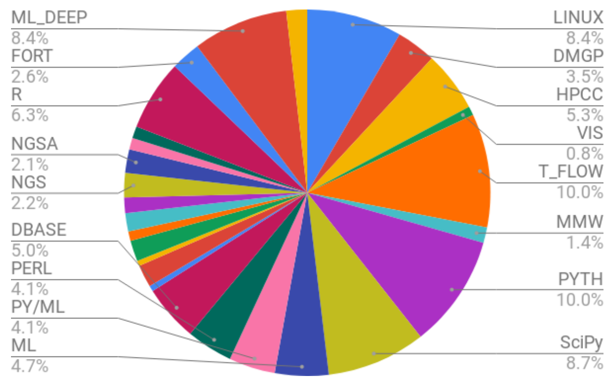

Figure 2. Distribution of participant registrations in HPRC short courses during Spring 2018. The figure displays combined attendance for both the in-person and WebEx sessions for each course. Topics include Fortran (FORT), R programming language (R), MATLAB (ML), Machine Learning and Deep Learning with MATLAB (ML_DEEP), Python programming language (PYTH), Linux classes and workshops, (LINUX), Databases (DBASE), Molecular Modeling Workshop (MMW) Scientific Python (SciPy), Python for MATLAB users (PY/ML), Perl (PERL), Next Generation Sequencing (NGS), NGS Assembly (NGSA), HPC Cluster usage (HPCC), Visualization portal (VIS), Data Management Practices (DMGP), TensorFlow (T_FLOW) and other topics.

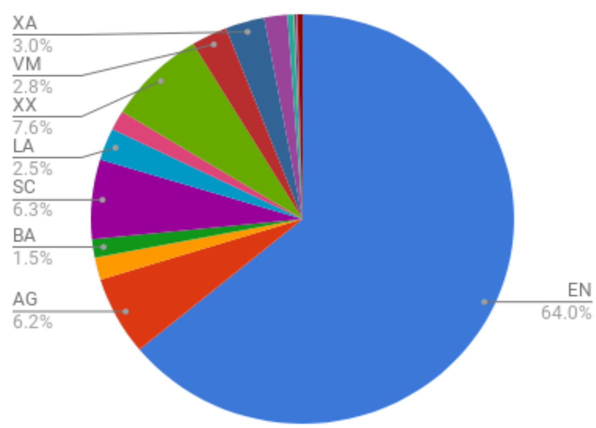

Figure 3. Distribution of registered participants in Spring 2018 across various TAMU colleges. The abbreviations used for the various colleges are Engineering (EN), Agriculture (AG), Business (BA), Science (SC), Liberal Arts (LA), Veterinary Medicine (VM), Other Entities including Industry (XX), and Other Academic Institutions (XA).

4.2 Developing Online Efforts

In Summer 2017, we observed that our data analysis bouquet of classes (Python and MATLAB) was routinely over-subscribed and the classrooms could no longer accommodate all interested participants. As we were on track to offer three courses in the following semester (Fall 2017), offering repeat courses for popular topics was not a viable option. To ensure that all researchers interested in taking these in-demand courses had an opportunity to benefit from them, we started offering HPRC short courses over WebEx in Fall 2017. WebEx participants participate in the same hands-on exercises as in-person attendees. To achieve this, we include a monitored discussion channels for online participants and have opened usage by migrating the training platform from HPRC clusters to Jupyter notebooks hosted on virtual machines. The online class platforms were further standardized in Spring 2018. We noticed that these sessions had an equal number of participants register as the in-person sessions. Providing these courses online via WebEx helped us reach out to a number of non-TAMU participants across the nation. While the majority of non-TAMU attendees participate using WebEx, individuals from local universities have also attended in-person sessions in College Station.

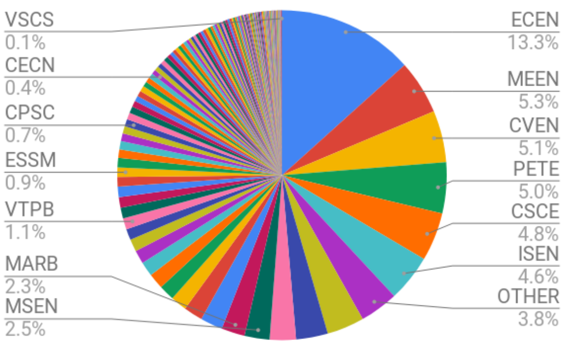

Figure 4. Distribution of registered participants across TAMU departments. Only departments with significant participation numbers are shown. Over 100 departments and institutions were served during this time frame. Percentage registrations from Electrical and Computer Engineering (ECEN), Mechanical Engineering (MEEN), Civil Engineering (CVEN), Petrochemical Engineering (PETE), Computer Engineering (CSCE), Industrial Science and Engineering (ISEN), Veterinary Small Animal Clinical Sciences (VSCS), Computer Science (CPSC), Ecosystem Science and Soil Management (ESSM), Veterinary Pathobiology (VTPB), Marine Biology (MARB), Materials Science and Engineering (MSEN) and non-TAMU units (Other) are shown.

4.3 Location and Participation

Broadening participation in computing is a core tenet of the HPRC training program. We have taken a number of steps to assist users from non-traditional fields of computing. Figure 3 describes the distribution of registered participants for Spring 2018 across TAMU colleges, other universities and industry. As TAMU is a predominantly engineering university, it is not surprising to note that the majority of participants in the TAMU HPRC short course program (64%) belong to the College of Engineering. While it is heartening to note the participation from the department of education, the limited participation from biology users is noticeably low for an agriculture-focused school. This "anomaly" in national trends is because a number of HPRC-themed bioinformatics courses are now taught by the Biology department in the "Digital Biology" course. An accompanying distribution of TAMU departments with the most registered researchers is provided in Figure 4. Unlike Figure 3, Figure 4 does indicate four biology departments with significant student participation.

TAMU HPRC short courses have been traditionally taught at a location that is close to most departments in the Colleges of Science and Engineering. TAMU main campus is divided into East and West Campus with buildings being miles apart. While the East campus has a stronger Engineering focus, the West Campus houses a number of the biology disciplines. It is possible to hypothesize that the location of our short courses may be a deterrent to participation from non-engineering disciplines. In an effort to rule out "location" as being a factor in the lack of short course attendance from, we hosted a number of short courses in the West Campus library that is located close to the Mays Business School, the College of Agriculture and the College of Veterinary Medicine in Fall 2017. While we observed a slight dip in participation from engineering departments, no similar uptick was observed from the departments located on West Campus. With the expanding online training platform further reducing the impact of location, all in-person courses were returned to their original locations. Indeed, overall that the trends observed in Figures 3 and 4 have been consistent over the last few semesters regardless of location, suggesting that location of these courses is not critical to their success. To assist our biology-oriented users, we moved our short course schedule from a mid-week schedule to an all-day Friday schedule so that our users would find it easier to park and would not have to travel back and forth across campus on multiple days. A longitudinal study is planned to test the hypothesis that the availability of courses over WebEx has reduced the importance of the location of our courses.

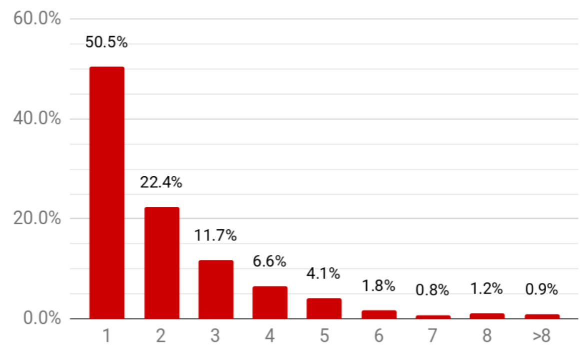

Figure 5. Student persistence profile for TAMU HPRC short courses in Spring 2018. Almost 50% of students enrolling in a HPRC short course returned for other courses. Participants enrolled in up to 14 short courses during the semester.

4.4 TIM and Student Persistence

In an effort to refine our offerings, we have investigated the participation profiles of our course attendees over Spring 2018. Figure 5 presents a student persistence profile that describes how many HPRC courses each participant registered for in Spring 2018. It is heartening to note that over 49% of participants re-registered for two or more courses. A significant number of these participants returned to take 4 or more courses over the Spring 2018 semester, with single participants registering for 13 and 14 courses as well. Though there was limited participation from students belonging to non-traditional fields of computational enquiry, those who did attend were most likely to register for a large number of all HPRC courses. This is not surprising considering that programming and HPC training is not commonly offered in these departments. These data present an emerging paradigm that students in disciplines that less computationally inclined are most likely to learn about computing from informal HPC short courses as opposed to formal courses. A similar longitudinal study about student attendance across multiple semesters is in the works. This data presents interesting insights into the design of a formal data sciences-oriented minor around these short courses.

The persistence information data indicates that a significant number of one-time participants register for some of the more popular HPRC courses (Machine Learning, CUDA, R, Python and MATLAB). As described above, TIM anticipates continued participation of students in foundational courses with a drop-off in participation as more advanced topics are covered. On the surface, the observed participation profile appears to violate the anticipated TIM learner progression, suggesting a reimagining of the instruction model. It is unclear whether technological interventions implemented by TAMU HPRC have allowed researchers to skip the foundation courses, or that immediate applicability of these materials to research forces students to cover the foundational materials on their own time. These data suggest that there is a need for the HPC-education community to develop TIM approaches for topics like Python and MATLAB. The TAMU HPRC short courses program is experimenting with a TIM model for Python. In Spring 2018, we offered a number of courses that introduced complexity in Python programming - Introductory Python, Applications of Scientific Python, TensorFlow, and Python for MATLAB users. In future iterations, we will include courses that cover topics in Parallel Applications of Python and the use of Pandas and Forecast Libraries.

It is interesting that the TAMU HPRC courses on R, MATLAB and Python continue to draw significant interest despite being supported by a number of formal efforts. In sharp contrast, our once popular "Introduction to CUDA" has since waned in popularity. This may be because, much like the case of Digital Biology formal classes, CUDA too is now being adopted in the formal Computer Science classroom. In contrast to CUDA, our classes with Artificial Intelligence or Machine Learning themes have been filled to capacity with students from Engineering. These participants represent a population of HPC users with different needs and are different from our typical user base. This became evident in our more traditional user-oriented courses. For example, attendees who took our "Scientific Python" course didn't appreciate the examples from Machine Learning training sets that were used as examples!

5. SUSTAINABILITY

The demonstrable need for TAMU HPRC short courses makes them inherently sustainable. They are offered free-of-charge to all participants on free-to-use software and machines. All course materials and notebooks are available for download free-of-charge from the TAMU HPRC website and we intend to release course recordings in the near future. The material from our short courses has been incorporated by courses currently taught at TAMU is currently being adopted by TAMU Galveston and Prairie View A&M University as well. The equipment for online WebEx broadcasts and video recordings is commercially available and may also be checked out from the TAMU libraries free-of-charge. Our National Science Foundation funded Cybertraining grant has provided us with the opportunity to develop a minor with a field of concentration in HPC as well. In addition, TAMU has submitted proposals that leverage the strengths of these short courses in search of federal dollars. As such, the approach toward institutionalizing the TAMU HPRC short course is likely to further strengthen the sustainability aspects.

6. EVALUATIONS & ASSESSMENTS

TAMU HPRC has worked with faculty in developing of "phase-gapped" evaluation strategies that help assess these programs. We are currently evaluating our initial designs in terms of their (1) connection to delivering key chemical and STEM concepts, (2) engagement and accessibility for students in a cyber learning context, and (3) support for instructors/peer leaders [Prince 2004, Parsons 2011]. TAMU HPRC currently collects data from two forms of evaluation: (1) a formative evaluation to assess the quality of project components, monitor project implementation, and provide ongoing feedback to the leadership team, and (2) a summative evaluation to examine the benefits to instructors and assess the impact of the project in reaching its stated goals. Both types of evaluation use a mixed method approach of qualitative and quantitative indicators. In the near future, we will also evaluate progress on the decided learning objectives, including our effectiveness in student-teacher engagement and learning.

We have traditionally relied on in-person interviews for feedback from the community. Registration and attendance data further help identify the effectiveness of our short courses. We rapidly came to the realization, however, that while these data demonstrated the demand for our courses on campus they didn't inform us about the quality of our courses. We experimented with collecting short surveys about the courses in Fall 2017. While we initially followed a model of mailing surveys electronically, the returns were extremely limited. Physical post-class surveys that required attendees to complete a questionnaire were implemented in the tail end of Fall 2017. Spring 2018 represents the first semester when evaluations were standard to each HPRC short courses. We currently follow a post-training evaluation model that includes in-person interviews and a free-format survey questionnaire. The survey focuses on course content and the participant's objectives. A free-format survey with open-ended questions was chosen over a Likert-scale style survey to ask open-ended questions and not constrain our participant's choices. Our typical surveys, while anonymous, provide participants with the opportunity to provide their email addresses if they wish to be contacted. Since TAMU HPRC courses are taught by CI professionals who volunteer to teach these topics, questions about the quality of the instructor that are typical in surveys on formal courses are not included. The Spring 2018 survey was a one-page document that includes questions such as [i] Did you attend this course for research, personal or class needs? [ii] What did you expect to learn from this course? [iii] Did the course meet your objectives? [iv] What did you like about the course? [v] What would you like us to do differently? [vi] What other courses would you like HPRC to offer? [vii] If you would like to subscribe to HPRC announcements please provide us with your email address. [viii] Please provide any additional comments below.

While the feedback from the surveys has helped refine our program, we face challenges in efficiently quantifying the collected data. Responses to the free-format surveys have provided us with data points for course success that we had not considered. Analyzing these surveys tends to be laborious and leaves room for ambiguity and personal interpretation. For these reasons, we will transition to a Likert-scale style survey approach in Fall 2018. This survey has been developed with feedback from TAMU faculty who focus on education. Some key points that we will be addressing in our future surveys are:

The open-ended questions in our survey will be:

1. How did you learn about this course?
2. Why did you register for this course?

3. Is this course related to your research or degree plan?
4. What are the difficulties that you faced in this course?
5. What do you think are the strengths of this course?
6. What specific content/concepts in the course were particularly challenging for you?
7. What specific content/concepts in the course were particularly easy for you?
8. How do you plan on using what you learn in this course?
9. Who would you recommend the course to?
10. Please add any additional comments below.

The Likert Scale (1-5 scale) questions will be:

1. How easy was the course for you?
2. How satisfied are you with the course?
3. How likely are you to take the course again?
4. How likely are you to recommend the course to others?

Our future surveys will provide us a rationale for why people are attending our classes. We will use these data to build a quantitative model for evaluating course success that goes beyond repeat attendance, and finally develop a profile of the kind of students or groups that are most likely to take our courses. In the future, we will partner with research groups on campus for an Internal Review Board (IRB) approved study to investigate whether attendees at our courses are meeting their desired learning objectives. Over the next few semesters, we will correlate data from assessments with course registration profiles to develop a teaching model for specific short courses. These steps are critical to shape the design of future HPC short courses type of efforts for users in non-traditional fields of computing.

7. CONCLUSIONS & LESSONS LEARNED

TAMU HPRC has implemented a number of interventions to drive a 300% growth in participation in HPRC short courses. In addition to the general need for data analytics in the scientific workflow, this growth may largely be attributed to our social and curricular interventions. Our data suggests that the influencing factors include offering courses on exciting topics, making the community aware of these courses, better student engagement by using active learning methods, avoiding policy bottle necks that curbed user participation, and finally by supporting our users with better documentation and support. The data from the TAMU HPRC short courses program supplies interesting insights on widely accepted models of the "tiered instruction" approach such as TIM. A longitudinal analysis of this data is further required to entirely understand these effects. While our current assessments have allowed us to refine our courses significantly, we will be utilizing stronger research-based methods in the near future. We will further standardize the active-learning segments of our courses so that all participants are guaranteed a similar experience in all of our courses. Over the coming semester we anticipate that developing ADA-compliant online courses will be our single-largest legal and administrative challenge. In our previous curriculum revisions, we have found that initial student resistance to new approaches can be overcome by good communication and persistence. Once students get accustomed to new approaches and expectations, they quickly regain their comfort level. We have utilized our student workers to serve as tutors and act as liaisons in resolving issues that arise. These steps have placed us on a firmer footing to develop this home-grown HPC effort into a certificate program.

8. SUPPORTING INFORMATION

All training materials developed by TAMU HPRC are available for download free-of-charge on the TAMU HPRC website. Please access the material at https://hprc.tamu.edu/training and send feedback about your adoption experience to help@hprc.tamu.edu.

9. ACKNOWLEDGMENTS

The authors would like to thank staff, student workers and researchers at Texas A&M HPRC, the Laboratory for Molecular Simulation, TexGen, Division of Research and Provost IT for supporting the HPRC short course program at Texas A&M University. Portions of this research were conducted on the Ada and Terra clusters, and virtual machines provided by TAMU HPRC. We gratefully acknowledge support from the NSF Abstract #1730695, CyberTraining: CIP: CiSE-ProS: Cyberinfrastructure Security Education for Professionals and Students. Special thanks to the instructors of each course: Dylan Rodriguez, Michael Dickens, Donald Rick McMullen, Lisa Perez, Mark Huang, Ping Luo, Jian Tao, Yang Liu, Marinus Pennings, Keith Jackson, Noushin Ghaffari, and Shichen Wang. We also gratefully acknowledge support from the Francis Dang, Mark Huang and Jack Perdue for maintaining the clusters and virtual machines used in these efforts.

10. REFERENCES

[1] Adams, C.M., and Pierce, R.L. " Teaching by tiering, " in Science and Children, vol. 41, no. 3, pp. 30-34, 2003

[2] Brennan, K., and Resnick, M., 2012 New frameworks for studying and assessing the development of computational thinking. In *Annual American Educational Research Association meeting* (Vancouver, BC, Canada)

[3] Lu, J. J. and Fletcher, G. H., 2009. Thinking about computational thinking. In *Proceedings of the 40th ACM Technical Symposium on Computer Science Education* SIGCSE '09, (New York, NY, USA), pp. 260–264, ACM.

[4] Parsons, J. and Taylor, L., 2011. Improving student engagement. Current issues in education, vol. 14, no. 1.

[5] Prince, M., 2004. Does active learning work? a review of the research. In *Journal of engineering education*, vol. 93, no. 3, pp. 223–231.

[6] National Science Foundation website describing the National Strategic Computing Initiative. Access at - https://www.nsf.gov/cise/nsci/

[7] Ontario Ministry of Education. (2005). *Education for All: The report of the expert panel on literacy and numeracy instruction for students with special education needs, Kindergarten to Grade 6.* Toronto, Ontario: Queen's Printer for Ontario. Access at: http://www.oafccd.com/documents/educationforall.pdf

[8] Ontario Ministry of Education. (2013). *Learning for all: A guide to effective assessment and instruction for all students, Kindergarten to Grade 12*. Toronto, ON: Queen's Printer for Ontario. Access at: http://www.edu.gov.on.ca/eng/general/elemsec/speced/LearningforAll2013.pdf

[9] Research on Learning in Formal and Informal Settings, National Science Foundation.

[10] Texas Regional Collaboratives, 2014. Building the Texas Computer Science Pipeline Strategic Recommendations for Success | theTRC.org

[11] Tomlinson, C.A, 1999. The Differentiated Classroom: Responding to the Needs of All Learners. Alexandria, Va.

[12] Wing, J. M., 2008. Computational thinking and thinking about computing. In *Philosophical transactions of the royal society of London A: mathematical, physical and engineering sciences*, vol. 366, no. 1881, pp. 3717–3725.

[13] Yadav, A., Zhou, N., Mayfield, C., Hambrusch, S., and Korb, J. T., 2011. Introducing computational thinking in education courses. In *Proceedings of the 42nd ACM technical symposium on Computer science education*, pp. 465–470, ACM, 2011.