Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025 ASEE
Paper ID #48109

# WIP: Formative Findings from the First Year Implementation of a Water and Data Science Workshop

**Dr. D. Matthew Boyer, Clemson University**

Dr. D. Matthew Boyer is a Research Associate Professor of Engineering & Science Education and an Educational Proposal Writer in the College of Engineering, Computing and Applied Sciences at Clemson University.

**Lukas Allen Bostick, Clemson University**
**Prof. Ibrahim Demir, The University of Iowa**
**Bijaya Adhikari**
**Krishna Panthi, Clemson University**
**Vidya Samadi, Clemson University**
**Mostafa Saberian, Clemson University**
**Carlos Erazo Ramirez, The University of Iowa**

# WIP: Formative Findings from the First Year Implementation of a Water and Data Science Workshop

## Introduction

The increasing complexity of global water challenges, such as resource management, climate resilience, and sustainability, demands interdisciplinary approaches that integrate advanced computational methods with domain-specific expertise [1]. With its capacity to process, analyze, and model large datasets, data science has become a transformative tool in water science and hydrology. By leveraging computational tools such as machine learning, hydrological modeling, and geospatial analytics, researchers can address pressing questions in water resource management more efficiently and effectively. Despite this potential, many professionals in the water science community lack access to the training necessary to harness these tools effectively, highlighting the critical need for cyber-training programs that bridge this skills gap.

The WaterSoftHack initiative, funded by the National Science Foundation CyberTraining program, is a multi-year effort to advance water science research through cybertraining and machine learning education. Anchored by the principles of inclusivity and collaboration, the project aims to democratize access to advanced cyberinfrastructure (CI) tools and workflows within the water science community. Key objectives include fostering workforce competence in CI prototyping and workflow development, spearheading innovation in sustainable water practices, and broadening the adoption of computational methods across the field. The program seeks to cultivate a new generation of interdisciplinary scientists equipped to address water-related challenges by targeting graduate and undergraduate students, early career researchers, and faculty. The capstone projects encourage participants to apply data science tools and concepts to real-world water science problems. This format fosters collaborative learning, enabling participants to leverage their diverse expertise while working under realistic constraints.

In this work-in-progress paper, we share an iteration of a longer design-based study as we assess the first year of the WaterSoftHack program to understand its impact on participants' educational and professional development. Using a mixed-methods approach, we analyze pre- and post-workshop survey data alongside qualitative insights from follow-up interviews. This analysis aims to identify the program's strengths and areas for improvement, providing actionable recommendations for future iterations. Building on two established evaluation frameworks, Kirkpatrick's [2] and CIPP [3], drawn from extension education research, this paper positions the WaterSoftHack program within a broader context of data-driven educational innovation focused on enabling educators to design, implement, and continuously improve educational initiatives.

Concerning our research questions and the objectives of this study, two key questions guide this work:

- How does the WaterSoftHack program enhance data science competencies within water science contexts?
- What challenges and opportunities do participants encounter as they integrate data science into their professional practices?

By addressing these questions, our objectives are to evaluate the program's effectiveness, identify areas for improvement, and contribute actionable insights to the field of data science education. The findings aim to inform iterative enhancements within the WaterSoftHack initiative and similar interdisciplinary educational programs, ensuring a sustained impact on participants' skills and broader educational practices.

## Literature Review

### *Data Science Education in Interdisciplinary Contexts*

Data science education thrives at the intersection of computational and domain-specific knowledge, particularly in water science. Research underscores the importance of interdisciplinary training to address complex, real-world problems. Manduca and Mogk [4] emphasize the value of integrating authentic research experiences and problem-solving into educational programs. The WaterSoftHack program embodies this approach by embedding data science training within the context of water science, targeting graduate students, postdoctoral researchers, and faculty members. This alignment ensures participants gain both technical proficiency and domain-specific insights.

### *Professional Development Workshops*

Workshops are pivotal in fostering continuous learning and skill development in fast-evolving fields. Desimone et al. [5] identify active learning, collaboration, and real-world relevance as key elements of effective professional development. The WaterSoftHack program integrates these principles by combining hands-on training with collaborative hackathon activities. Participants are challenged to apply newly acquired skills to water-related projects, ensuring that the training directly translates to practical applications.

### *Evaluation Frameworks*

Given the similarities in our study with extension education research, we use two models, Kirkpatrick's and CIPP, from extension education to inform our perspective on ensuring the program is effective and responsive to community needs. Kirkpatrick's evaluation model assesses training effectiveness across four levels, i.e., Reaction, Learning, Behavior, and Results [2]. The model is used in extension education to measure participant satisfaction and the translation of learning into real-world impact, much like the goals of the WaterSoftHack program. The CIPP model evaluates programs through Context, Input, Process, and Product dimensions [3]. This holistic approach helps ensure that program design, implementation, and outcomes are well-aligned with local needs. In addition to these models, recent work by Sajja et al. [6] on integrating Generative AI in hackathons further illustrates innovative evaluation techniques that enhance learning and program assessment.

### *Program Evaluation Through Surveys and Interviews*

Combining surveys and interviews offers a robust framework for evaluating program outcomes. Surveys provide quantitative metrics, such as participant satisfaction and confidence levels, while interviews deliver rich, qualitative insights into individual experiences. Creswell, Plano Clark, and others [7] advocate for a mixed-methods approach to program evaluation to capture general trends and nuanced feedback. In the WaterSoftHack program, pre- and post-workshop surveys, complemented by follow-up interviews, enable a comprehensive assessment of participant outcomes and areas for improvement.

### *Hackathon-Style Training Methods*

Hackathons have emerged as innovative platforms for skill development, fostering collaboration, problem-solving, and rapid prototyping [6]. Research by Briscoe and Mulligan [8] highlights the effectiveness of hackathons in promoting interdisciplinary teamwork and accelerated learning. The WaterSoftHack program leverages this format by tasking participants with tackling real-world water science challenges in diverse, collaborative teams. This approach enhances technical skills and mirrors professional environments' dynamic, problem-focused nature.

## Program Description

### *Workshop Design*

The workshop integrated water and data sciences in a three-week virtual training program designed to develop participants' data science skills within the context of water science and combined asynchronous preparatory materials, live instructional sessions, and team-based problem-solving activities. Project team members introduced participants to tools and techniques such as data retrieval, programming with JavaScript, and hydrological modeling, culminating in a hackathon-style capstone project. The design for content delivery emphasizes interdisciplinary collaboration, enabling participants to work in teams with diverse expertise. Instructional materials included detailed tutorials, datasets for hands-on practice, and video lectures to support self-paced learning. The hackathon challenged teams to apply their skills to real-world problems, fostering innovation and teamwork under time constraints.

### *Participant Profile*

The workshop attracted graduate and undergraduate students, early career researchers and faculty members from various disciplines, including civil engineering, environmental sciences, and computer science. Participants represented a mix of experience levels, ranging from beginners to those with prior exposure to data science methods. Financial support to participants was provided through NSF CyberTraining grants, ensuring accessibility for diverse attendees.

### *Instructional and Material Design*

The preparatory materials provided a foundation for the workshop, covering fundamental programming concepts and data science techniques. Live sessions focused on hands-on activities, guided by experienced instructors with water science and computational methods expertise. Collaboration tools, including virtual whiteboards and shared coding environments, enhanced team interactions and project execution.

## Research Methods

### *Design-Based Research Approach*

The workshop’s development and evaluation followed a design-based research methodology, emphasizing iterative refinement based on participant feedback. This approach allowed for continuous improvements to instructional content and delivery methods, informed by data collected during the first year of implementation.

### *Data Collection*

The educational research team members collected data through pre- and post-workshop surveys and follow-up participant interviews (e.g., Figure 1). The surveys included Likert-scale and open-ended questions to assess satisfaction, skill development, and confidence in applying data

science methods. Follow-up interviews were semi-structured, focusing on participants' experiences, perceived challenges, and suggestions for improvement.

### *Data Analysis*

We used a mixed-methods approach to analyze the data. Quantitative survey responses were summarized using descriptive statistics while thematically coding qualitative feedback from open-ended survey questions and interviews. This combined analysis provided a comprehensive understanding of participant outcomes and informed recommendations for program refinement.

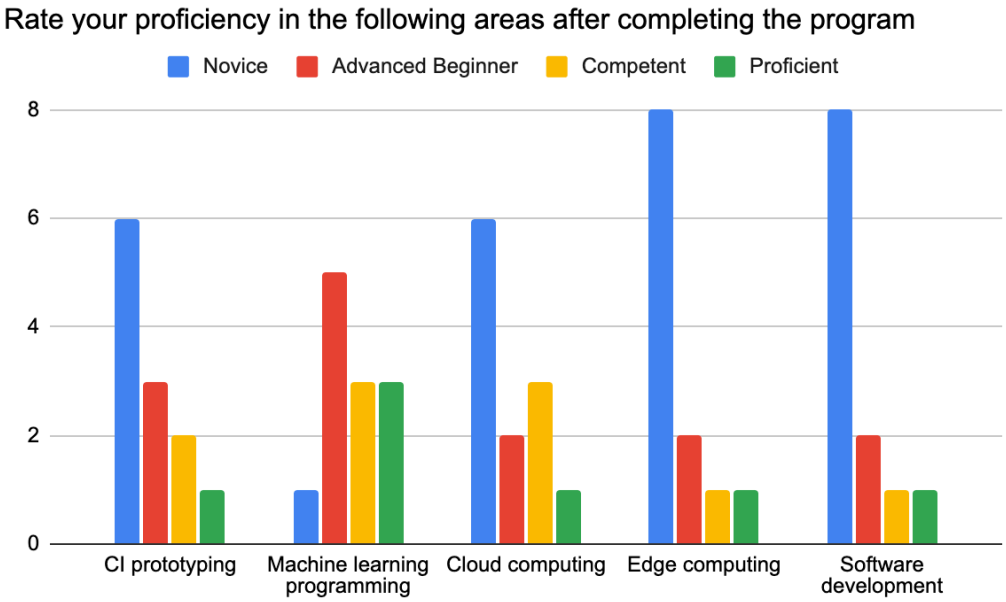


*Figure 1. Example Survey Question and Response*

### *Limitations*

The primary limitations of this study include the small sample size and potential selection bias in the interviews, as only a small subset of participants volunteered for follow-up discussions, i.e., two of fifteen invited. Given this low response rate to interviews, we used interview findings as ways to illuminate rather than validate survey findings across participants. Additionally, the virtual format may have influenced participant engagement and feedback, limiting generalizability to in-person workshops.

## Findings

### *Survey Results*

The survey responses revealed moderate satisfaction with the WaterSoftHack event. On a five-point Likert scale, with 5.0 being the highest rating, participants rated overall satisfaction as 3.8 on average, indicating general approval with room for improvement. Ratings for the workshop's effectiveness in meeting learning and professional goals were slightly lower, averaging 3.5. Participants noted the instructional materials as one of the workshop's strengths, with a mean rating of 4.2. Confidence in applying newly acquired skills post-workshop varied significantly, with ratings ranging from 2 (not confident) to 5 (very confident), averaging 3.4.

Regarding specific skill areas, participants indicated the most growth in understanding data retrieval processes and basic programming techniques. However, proficiency ratings for tools such as JavaScript and advanced modeling techniques suggested these areas were more challenging. Open-ended survey responses highlighted time constraints, uneven prior knowledge among participants, and the compressed format of the hackathon as areas requiring attention.

### *Interview Results*

Follow-up interviews provided qualitative depth to the survey findings, offering nuanced perspectives on participants' experiences. Both interviewees appreciated the preparatory materials, describing them as essential for building foundational knowledge. However, they also noted limited engagement with these materials due to time constraints before the workshop. The interviews underscored variability in participants' starting points, with one interviewee entering the program as a complete beginner and the other having prior exposure to data science concepts.

The hackathon emerged as a focal point of discussion, with participants praising its collaborative nature and the opportunity to engage with real-world datasets. Both interviewees suggested structural adjustments, such as an earlier introduction of project topics and additional time for team preparation. They also described challenges related to the virtual format, noting its benefits for accessibility but expressing a preference for in-person interaction for certain aspects of the workshop, such as team collaboration.

Participants highlighted networking opportunities as a key benefit across findings, with interviewees emphasizing the value of interdisciplinary collaboration and the potential for future partnerships. Both participants also discussed their plans to integrate workshop content into their professional workflows. However, they acknowledged that doing so would require additional practice and reinforcement of the skills introduced during the workshop.

## Integrating Interview and Survey Insights

In this section, we synthesize insights from the survey responses and follow-up interviews to present a nuanced understanding of participant experiences in the WaterSoftHack event. By examining both datasets, we aim to highlight areas of alignment and divergence to inform the iterative improvement of the workshop design.

### *Participant Satisfaction and Program Effectiveness*

Survey data revealed that participants were moderately satisfied with the workshop, with ratings generally clustered around the midpoint of the satisfaction and program effectiveness scales. The interviews corroborated this trend, as participants appreciated the workshop's foundational content and relevance to their academic and professional goals. Both interviewees described the program as a stepping stone toward developing proficiency in data science methods, with one participant explicitly noting that the program helped overcome their initial apprehension about programming.

Interviews provided additional depth to these findings. For instance, while survey responses aggregated satisfaction broadly, interviewees shared specific aspects they found valuable, such as the preparatory materials and the opportunity to work with real-world datasets. One participant noted that while the workshop effectively laid the groundwork for future applications, the short

timeframe constrained their ability to engage with the material thoroughly. This insight highlights the need for more detailed survey questions that capture time-related challenges.

***Preparation Materials and Skill Development***

Both the survey and interviews underscored the high value of the preparatory materials. Participants rated these resources highly effective, and interviewees elaborated on their impact. One interviewee emphasized that the preparatory materials provided crucial foundational knowledge, enabling them to approach the workshop confidently. However, both interviewees expressed regret over not having dedicated more time to these materials before the start of the workshop. This feedback suggests the potential benefit of requiring or incentivizing pre-workshop engagement to ensure participants are adequately prepared to participate fully during the synchronous event.

Survey responses indicated moderate confidence in participants' ability to apply the skills they developed during the workshop. Similarly, interviewees described increased confidence but acknowledged that it was primarily related to understanding processes and tools rather than mastering them. For example, one participant noted that the program introduced them to essential programming packages and data retrieval processes, which they plan to incorporate into their dissertation work in the future.

***Collaboration and Networking***

Survey respondents valued the networking opportunities facilitated by the workshop, and interviews offered richer narratives around this theme. Both interviewees highlighted the diversity of expertise within their teams as a significant strength. One participant shared that collaborating with peers from different disciplines broadened their perspectives and validated their approach to research. The other interviewee emphasized that these connections could lead to future collaborations, particularly in areas where their expertise was limited.

Despite these benefits, the interviews also revealed challenges related to collaboration. Beginners sometimes felt hesitant to contribute fully, fearing they might hinder the progress of more advanced team members. While this dynamic did not emerge in the survey data, it underscores the importance of effectively designing team structures and support mechanisms that balance diverse skill levels.

***Hackathon and Workshop Structure***

Both survey and interview data highlighted the hackathon as a pivotal component of the workshop. Participants generally rated the hackathon as moderately effective, with interviewees elaborating on its strengths and areas for improvement. One participant suggested that introducing project topics earlier in the workshop would allow teams more time to explore and prepare, potentially enhancing the quality of their outcomes. Another noted that the condensed timeframe limited their ability to fully engage with the project, particularly as a beginner.

While survey responses did not explicitly address the virtual format, the interviews provided contrasting perspectives. One interviewee found the virtual format convenient and well-suited to the workshop's content, while the other preferred in-person interactions, citing the potential for

deeper engagement. This divergence highlights an opportunity to explore hybrid models that combine the accessibility of virtual sessions with the benefits of face-to-face collaboration.

***Recommendations for Improvement***

The integration of survey and interview findings suggests several actionable recommendations for enhancing the WaterSoftHack event:

- Extended preparation period: encourage participants to engage with preparatory materials through mandatory assignments or pre-workshop sessions.
- Team structuring for diverse expertise: design team compositions to balance varying skill levels and provide clear guidance on roles to empower all participants to contribute effectively.
- Earlier introduction of project topics: introduce hackathon themes and project options at the start of the workshop to maximize preparation and engagement.
- Adjustments to workshop format: consider hybrid models that leverage the strengths of both virtual and in-person formats.
- Incremental skill development: incorporate follow-up sessions or resources between workshops to reinforce learning and maintain participant momentum.

**Discussion**

The formative findings from the first year of the WaterSoftHack program highlight the program's initial successes in addressing the skills gap at the intersection of data and water sciences. Participant feedback suggests alignment with the program's core objectives of fostering interdisciplinary collaboration and developing computational competencies in water science contexts. Survey results indicate moderate satisfaction and skill development, complemented by interview insights emphasizing the program's value in building foundational knowledge and professional networks.

One noteworthy takeaway is the role of the hackathon in promoting applied learning. Participants valued its collaborative and problem-oriented nature, which mirrored real-world scenarios. However, challenges such as the short timeframe and uneven team dynamics suggest opportunities for improvement. Extending preparation periods, refining team structures, and introducing project themes earlier could enhance participants' experiences in future iterations.

The mixed-methods evaluation further highlights the complementary strengths of surveys and interviews in capturing participant outcomes. While surveys provided quantitative measures of satisfaction and confidence, interviews offered richer, qualitative insights into individual experiences and program impact. This dual approach enabled a more nuanced understanding of the program's strengths and areas for improvement. These findings underscore the importance of iterative refinement in design-based research. Adjustments informed by this evaluation, such as exploring hybrid workshop formats and incremental skill-building opportunities, will ensure that the program continues to meet the evolving needs of its participants.

The findings from the WaterSoftHack program offer valuable insights that can inform similar educational initiatives. The positive outcomes observed, such as enhanced teamwork, creativity, and real-world problem-solving skills, align with results from other hackathon-based learning environments, underscoring the efficacy of such approaches in fostering practical competencies

[9]. We identified challenges when implementing our program, including resource constraints and managing cultural diversity, mirroring those reported in comparable settings, and highlighting common areas needing attention to successfully implement similar experiences [9]. Using Kirkpatrick's and the CIPP frameworks as evaluation models [2], [3], and [10], our approach offers a structured methodology that can be adapted to assess and enhance the effectiveness of diverse educational programs, ensuring they are contextually relevant and outcome-focused given attention to essential differences in implementation environments.

Examining the findings through the lenses of Kirkpatrick's and CIPP models provides a multifaceted view of the program's impact. From Kirkpatrick's perspective, survey data indicates moderate participant satisfaction (Reaction) and some improvements in foundational skills (Learning), though the challenges reported in applying advanced techniques suggest opportunities for enhancing behavioral change (Behavior) and overall outcomes (Results). The CIPP framework reveals that the program's context, i.e., addressing urgent water science challenges, aligns well with local and national needs, while the evaluation of inputs and processes underscores both the strengths of the workshop's design and areas requiring refinement, such as team dynamics and time allocation. Together, these frameworks help to inform the successes of WaterSoftHack and highlight clear pathways for iterative improvement, ensuring the program remains responsive and impactful in the evolving landscape of data-driven education.

**Conclusions**

The first year of the WaterSoftHack program demonstrated its effectiveness as a model for interdisciplinary data science education, particularly within water science. Participants reported notable gains in foundational data science skills and valuable insights into applying computational methods to water-related challenges. However, the program's intensive schedule and virtual format presented difficulties for some, indicating areas for enhancement. Recommendations derived from this evaluation include bolstering preparatory resources, fostering improved team dynamics, and exploring hybrid learning models that combine virtual and in-person interactions.

The WaterSoftHack program contributes to the broader landscape of data science education and professional development by bridging computational techniques with water science. As the program progresses into its next phase in Summer 2025, these evidence-based insights will guide ongoing refinements, ensuring the initiative democratizes access to advanced computational tools and cultivates an inclusive community of water science professionals.

**Acknowledgment**

National Science Foundation Award 2320979 supports this work. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

**References**

[1] C. Ramirez, Y. Sermet, and I. Demir, "HydroLang Markup Language: Community-driven web components for hydrological analyses," *J. Hydroinform.*, vol. 25, no. 4, pp. 1171–1187, Aug. 2023.

[2] D. Kirkpatrick and J. Kirkpatrick, *Evaluating Training Programs: The Four Levels*. Oakland, CA, USA: Berrett-Koehler Publishers, 2006.

[3] D. L. Stufflebeam, "The CIPP model for evaluation," in *International Handbook of Educational Evaluation*, T. Kellaghan and D. L. Stufflebeam, Eds. Dordrecht, The Netherlands: Springer, 2003, pp. 31–62. doi: 10.1007/978-94-010-0309-4_4.

[4] C. A. Manduca and D. W. Mogk, Eds., *Earth and Mind: How Geologists Think and Learn about the Earth*. Boulder, CO, USA: Geological Society of America, 2006.

[5] L. M. Desimone, N. Bell, A. Lentz, K. L. Hill, and L. Marianno, "A holistic examination of how professional learning and curriculum relate to ambitious and culturally relevant instruction and student engagement," *AERA Open*, vol. 11, Art. no. 23328584241310428, Mar. 2024. doi: 10.1177/23328584241310428. (Note: Year corrected based on DOI publication date).

[6] R. Sajja, C. E. Ramirez, Z. Li, B. Z. Demiray, Y. Sermet, and I. Demir, "Integrating generative AI in hackathons: Opportunities, challenges, and educational implications," *Big Data Cogn. Comput.*, vol. 8, no. 12, Art. no. 188, Dec. 2024. doi: 10.3390/bdcc8120188.

[7] J. W. Creswell and V. L. P. Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Thousand Oaks, CA, USA: SAGE Publications, 2017.

[8] G. Briscoe and C. Mulligan, "Digital innovation: The hackathon phenomenon," Creativeworks London, London, U.K., Working Paper 6, 2014. [Online].

[9] T. Z. K. Oyetade and A. Harmse, "Evaluation of the impact of hackathons in education," *Cogent Education*, vol. 11, no. 1, Art. no. 2392420, Aug. 2024. doi: 10.1080/2331186X.2024.2392420.

[10] R. Gandomkar, "Comparing Kirkpatrick's original and new model with CIPP evaluation model," *J. Adv. Med. Educ. Prof.*, vol. 6, no. 2, pp. 94–97, Apr. 2018.