

# 1st Science Understanding through Data Science Conference

## L<sup>A</sup>T<sub>E</sub>X Submission Template

Jake Lee<sup>\*1</sup>, Sam Berndt<sup>1</sup>, Lukas Mandrake<sup>1</sup>, Erika Podest<sup>1</sup>, and Ryan McGranaghan<sup>1</sup>

<sup>1</sup>Jet Propulsion Laboratory, California Institute of Technology, CA, USA

## 1 Introduction

**(Word Limit: 500)** We are seeking abstracts that are aligned with one or more of the following categories. Please directly address both the technical approach and the nature of the collaboration between physical and data scientists in your abstract [1]. Note that data science includes, but is not limited to, machine learning, low-order parametric models, and uncertainty quantification.

- Collaborative works involving both physical scientists and data scientists
- Physical scientists demonstrating modern data science techniques with science outcomes
- Novel data science approaches motivated by physical science applications
- Works with emphasis on explainability, transparency, causality, or quantifiable uncertainty
- Novel physical science investigations enabled by data science approaches
- Operational data science-based implementations in science data pipelines (a.k.a. Science MLOps)
- Quantifying and assessing the effectiveness of physical science-data science collaboration
- Institutional strategies for organization of SUDS-like efforts, disciplines, and evaluation
- Visions for the future of physical science-data science collaboration

The SUDS mission statement is “To increase the speed, depth, and rigor of scientific return by revealing new connections through data science.” The goal is to produce more than “an accurate prediction model,” but to facilitate new insight and scientific understanding. In this section, please describe the research domain: What’s the science question, why is it interesting, and what will be the impact of addressing it? What is the status quo, and why isn’t it working? What aspect of the question/problem seems appropriate for data science, and why?

## 2 Approach

Describe the data science by translating the science challenge into a clear data science objective and appropriate success metric. Discuss the data sets (size, type, content), the data science approach, what other approaches

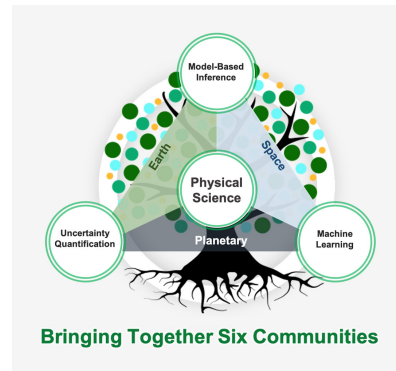


Figure 1: Only one figure is allowed in the abstract. The caption should be descriptive and stand alone. The figure should be placed at the top of the second column. Caption text **does** count towards the word limit.

were considered, and any reusable frameworks or codebases generated that others can benefit from. Discuss the compute resources and costs if relevant.

## 3 Results and Conclusion

Discuss the findings: what did the data science find, and which parts were obvious (confirmatory) vs. insightful (surprising)? How did these findings compare to traditional methods of solution and/or existing community approaches? What are the challenges this approach brings, and how can others overcome it? What other science problems could benefit from a similar approach?

## 4 Collaboration

If relevant, speak to the collaboration between the data and the physical scientist: what were the challenges and how were they overcome? Relate these findings to the **Community of Practice**: lessons learned, recommendations, future needs, institutional support that could have made things easier, etc.

### Acknowledgements

Please include your funding acknowledgement here. This text **does not** count towards the word limit.

<sup>\*</sup>Organizing Committee Contact: sudsconf@jpl.nasa.gov

## References

- [1] Umaa Rebbapragada, Serina Diniega, Gary B Doran, Steven Lu, Mark Wronkiewicz, and Jacob Widmer. Forming a science and data science collaboration to better understand the global martian frost cycle. *AGU23*, 2023.