FIELD EVALUATION OF ROBOTIC INFORMATION GATHERING ALGORITHMS WITH EXPERT EARTH AND PLANETARY SCIENTISTS.

Z. I. Lee[1], I. Brunton[2], R. C. Ewing[3], K. R. Fisher[2], D. J. Jerolmack[4], N. A. Jones[5], D. E. Koditschek[4], M. Nachan[3], F. Qian[6], F. Rivera-Hernandez[7], J. Ruck[4], T. Shipley[8], S. Thompson[7], C. G. Wilson[1]. [1]Oregon State University (leeza@oregonstate.edu), [2]NASA Johnson Space Center, [3]Texas A&M University, [4]University of Pennsylvania, [5]Northern Arizona University, [6]University of Southern California, [7]Georgia Institute of Technology, [8]Temple University

Introduction: Robots are increasingly used in scientific data collection on Earth and other planets for their ability to provide high-accuracy, multi-sensor data at high spatiotemporal resolutions. As robotic capabilities have advanced, roboticists have put significant effort into developing robotic information gathering algorithms [1,2,3]. Despite advancements, these algorithms are not widely used in science missions, with most information gathering decisions still being made by mission scientists. Here, we present early efforts to understand the reasons for slow algorithm uptake, by characterizing scientists' perceptions of the success of robotic information gathering algorithms during a field campaign. We completed case studies of four scientists to evaluate their satisfaction with two 'off-the-shelf' robotic information gathering algorithms during a lunar analogue mission on Mt. Hood in Oregon, USA.

Our case studies evaluated scientist satisfaction with two types of algorithms; a 'greedy' algorithm, and an algorithm based on Monte-Carlo Tree Search (MCTS) methods. Greedy algorithms are algorithms which take the most immediately profitable action [4]. MCTS algorithms, in comparison, repeatedly simulate possible sets of actions in search of a more optimal solution, with a bias towards sets of actions that seem promising [5]. We chose to use a MCTS algorithm as they are used in various robotic information gathering tasks [6,7], and chose to use a greedy algorithm as they are a standard baseline for algorithm comparison experiments. For our study, we developed a simple algorithm which matches concepts laid out in [4], while our MCTS algorithm was based on work from [8].

Methods: The algorithms were assessed during an analogue mission deploying a legged robot for data collection on Mt. Hood (See Fig. 1A). The overarching goal of the mission was to investigate how ice content alters regolith strength, as measured via robot leg-surface interactions. The science team did not have strong a-priori beliefs about ice-strength relationships, more interested in exploring the space of possible relationships than exploiting one particular hypothesis. Upon arrival at the field location, a sampling area for testing the algorithms was identified by a scientist who did not take part in the evaluation (See Fig. 1B).

Both of the algorithms evaluated used elevation and distance from the edge of an ice patch (referred to as ice-regolith boundary distance) as factors for selecting sampling locations. Each algorithm generated a sampling plan consisting of 12 valid sampling locations within the identified sampling area (See Fig. 1C). Locations were considered valid if they were within the identified sampling area and were within 2.5 meters of an ice-regolith boundary. The greedy algorithm divided the sampling area into smaller spatial areas, determined how much variation in elevation and ice-regolith boundary distance within them and selected two points from the six areas with the highest variation. The MCTS algorithm took a subset of 200 random sampling locations and repeatedly simulated combinations of 12 points to find a combination of points which maximizes average distance between normalized elevation values and normalized ice-regolith boundary values.

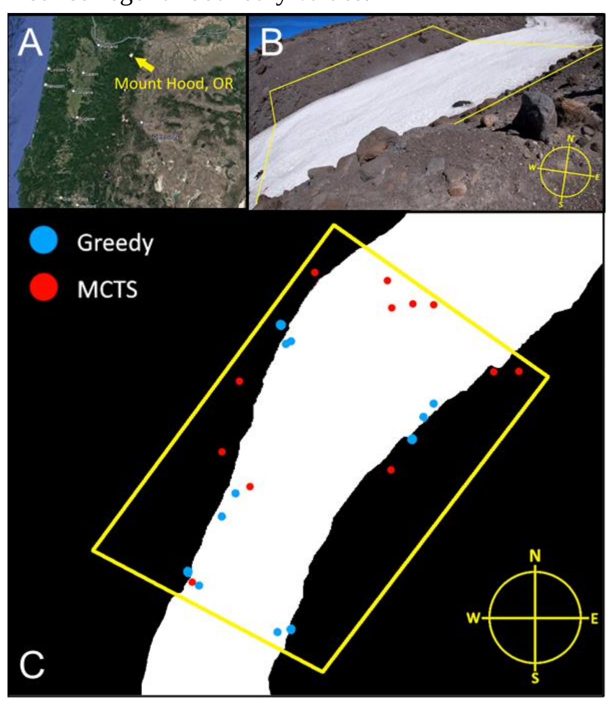

Figure 1. *(A) Map of Mt. Hood, (B) Photo of sampling area outlined in yellow, (C) Top-down representation of sampling area, white is ice patch, black is regolith, algorithmically selected sampling locations shown.*

Four members of the science team evaluated the information gathering algorithms during the field campaign. All participants were PhD students in Earth or planetary sciences from different eminent R1 institutions across the USA, and from different intellectual backgrounds and research traditions, including geophysics, planetary geology, and hydrogeology.

All case study participants were presented with the algorithmically generated data collection plans in the same order (MCTS first, Greedy second) and asked the same sequence of questions about both plans. Participants were asked to rate their satisfaction with the plan on a scale from one (least satisfied) to five (most satisfied) based on (i) coverage of the sampling area space, defined as the spread of locations in the latitude and longitude dimensions, and (ii) coverage of the ice-regolith gradient, defined as the spread of locations in the signed distance value from a location to the nearest ice-regolith boundary. Participants were also asked during a semi-structured interview to report any notable gaps and redundancies in the sampling plans, and to reflect on what each plan did well/poorly.

Results: Participants' reported satisfaction with algorithmically generated sampling plan was comparable across algorithm type (see Fig 2). Satisfaction with coverage of the sampling area and ice-regolith gradient was moderate for both the MCTS and greedy algorithms, with mean values between three and four on a five-point scale. Participants' responses during the semi-structured interview revealed the rationale behind the moderate satisfaction scores.

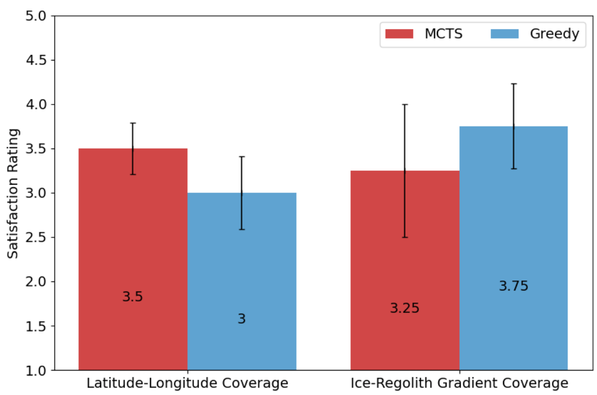

Figure 2. *Average satisfaction scores with +/- 1 standard error*

All participants reported that the spread of locations in physical space was less than ideal due to gaps and redundancies. Participants were in high agreement about areas of redundancy. All four participants identified the same pair of NW on-ice points in the greedy algorithm sampling plan as redundant, and three participants identified the same cluster of NE on-ice points in the MCTS algorithm as redundant. There was also convergence amongst participants about sampling plan gaps. Two participants identified the same gaps in the SE and NE for the greedy algorithm, and the SE for the MCTS algorithm. The other two participants, while not identifying specific gaps, reported the spread of sampling locations could be better and more systematic.

Participants varied in their perception of how successfully the sampling plans covered the ice-regolith gradient. In the interview, three participants reported that the sampling plans tracked the ice-regolith boundary well, and two of these participants also ranked high satisfaction with ice-regolith coverage. The other two participants, who ranked lower satisfaction with ice-regolith coverage, reported that there were insufficient sampling locations away from the ice-regolith boundary and an uneven balance of sampling location on ice versus regolith.

All participants reported wanting more information to form their judgments about the sampling plans. Two of the participants said they would ideally like to see the data collected prior to judging the plan. The other two participants said they would like more information on how the algorithms were making decisions. Interestingly, in the absence of being provided with explanations of algorithm choices, participants tended to ascribe more sophisticated decision making to the algorithms than was the reality. For example, a participant suggested during the interview that the algorithms had recognized and captured areas with different ground features, but this was not a known factor to the algorithms.

Conclusions: The results show that 'off-the-shelf' robotic information gathering algorithms generate sampling plans that are moderately satisfactory to scientists. Scientists were in high agreement about the presence of redundancies and gaps in sampling points in physical space. A subset of scientists also reported poor coverage of the ice-regolith gradient space. Therefore, satisfaction with algorithmic sampling plans could likely be improved in the future by incorporating scientists' rules-of-thumb for how sampling locations should be distributed in physical and gradient space. Future work should also consider how scientists' satisfaction with sampling plans is impacted by seeing the data collected, and/or by receiving explanations of algorithmic decisions.

Acknowledgments: This project is funded by the NASA PSTAR Program, grant 80NSSC22K1313. Special thanks to Ian Rankin for his assistance with coding.

References: [1] Shkurti, F., et al. (2012) *IROS*, 1747-1753. [2] Nilsson, P., et al. (2018) *RSS*, 1-9. [3] McMahon, J., et al. (2022) *AERO*, 1-11. [4] Vince, A. (2002). *Discrete Applied Mathematics*, 121(1-3), 247-260. [5] Fu, M. C. (2018) *WSC*. 222-236. [6] Guerrero, E., et al. (2021) *IEEE Access*, *9*, 136487-136506. [7] Nguyen, J. L., et al. (2014) *ICRA*, 6573-6578. [8] Rankin, I. C., et al. (2021) *ICRA*, 4882-4888.