AUSTRALIAN ROVER CHALLENGE

Founded by the University of Adelaide

AUSTRALIAN ROVER CHALLENGE 2025

SCORE REPORT

VERSION 1.0 09 JULY 2025

CATERPILLAR®

ispace

Table of Contents

Part I.

Summary and Overview

1. Judge Comments

Overall teams made significant improvements from last year (2024), however fundamental issues related to basic mobility and adherance to Rules & Regulations persist. For context, this document shall be read in conjunction with the 2026 installment of the ARCh Rules & Regulations document.

- New (first time attending ARCh) teams continue to struggle with basic functionality and mobility at ARCh. We **re-iterate** the importance of testing and readiness checks in the lead-up to ARCh, and the importance of reducing technical scope. As clearly stated in **Rule 4.5** in the Rules & Regulations, specifically **Rule 4.5.5**, a rover that can turn on, move and see is more important than a rover that can attempt all activities with inconsistent or little success.

- As demonstrated this year, the technical expertise of mature teams is sufficient whereby scores of 100% were achieved, indicating that upward difficulty adjustments are required. This requires a balanced approach that still considers new teams, who continue to struggle with basic functionality.

- Teams adhered to new radiofrequency (RF) communication modifications to **Rule 3.10.4** which resulted in improved performance and reliability for all teams. Although some issues persist that are often beyond the scope or control of the judging commitee, as stated in the new dispute resolution process described in **Rule 2.12**, we request that all teams record and attempt to quantify any issues, such as RF communications, to assist judges in continuing to improve the challenge for all competitors.

- It is strongly advised that all teams recognize new mandatory requirements to log power (**Rule 3.15.2**) during each task, and that these metrics along with the size and weight of the system(s) will be evaluated when calculating the Efficiency Multiplier (**Appendix D**) - a new process for rewarding good engineering practices. In short, a simple and lightweight rover that performs a task efficiently and effectively will receive more points than another rover that performs the same task with identical success, but with a heavier, inefficient system.

- Overall, the current level of difficulty for each task is considered appropriately challenging. Teams that optimise time allocation and the sequencing of activities during the task window, develop more robust and failure-resistant subsystems (to reduce time lost to malfunctions), and invest in operator training and effective intra-team communication—particularly for the base station and ground crews during high-stress scenarios like sub-system failure—will be well-positioned to achieve more points in these activities in the future.

- We strongly recommend teams improve their system by building on proven heritage using a near-identical platform or sub-systems, rather than a full redesign. Teams that ignored this in the past often performed worse—or failed to participate—while facing added financial, scheduling, and mental strain. This is not to say innovation and improvements should not be made, simply that the leadership team should, account for timeline and progress delays, budget, procurement, workshop, facility access and most importantly, team capability.

For Figures 2.1, 3.1, 4.1 and 5.1, the 'pass rate' is the percentage of teams that achieved any points for each sub-activity. Here, this is interpreted as the participation rate, which refelcts how many teams were able to participate in specific activities and represents a metric for difficulty.

See Figure 3 3 in the Appendix for individualized radar charts for each team's overall performance, with the average score for each deliverable and task overlaid as dashed green line.

1.1. Overall Summary

Table 1.1.: Team scores by task and total score

Team	CDR (/30)	SAR (/70)	Post-Landing (/100)	Space Resources (/100)	Excavation & Construc-tion (/100)	Mapping & Autonomy (/100)	Total
UoA	23.75	52.70	20.00	29.00	70.00	13.00	208.45
RMIT	19.40	49.23	30.00	43.80	33.00	15.00	190.43
Monash	26.78	62.00	65.00	68.00	92.00	45.00	358.78
UniMelb	23.12	54.40	58.50	25.50	44.20	9.00	214.72
UTS	21.18	50.00	4.50	42.40	26.00	9.00	153.07
UQ	22.28	55.33	85.00	83.00	100.00	33.00	378.60
Scorpio	20.28	57.03	55.00	71.00	56.00	33.00	292.30
QUT	23.15	55.10	70.00	28.00	65.00	29.00	270.25
UNSW-RAS	22.68	59.63	70.00	30.00	71.80	35.00	289.10
UoW	20.18	48.98	50.00	25.00	56.20	46.00	246.35
AGH	21.58	54.00	90.00	76.00	100.00	80.00	421.58
Legendary	16.44	59.05	90.00	27.40	89.00	37.40	319.29
Bluesat	22.08	49.93	0.00	11.00	9.00	0.00	92.01
Deakin	17.80	48.43	27.00	26.00	0.00	0.00	119.23
UWA	15.53	43.43	4.00	16.00	0.00	12.00	90.95
Minimum	15.52	43.42	0.00	11.00	0.00	0.00	90.95
Average	21.08	53.28	47.93	40.14	54.15	26.43	243.01
Maximum	26.78	62.00	90.00	83.00	100.00	80.00	421.58

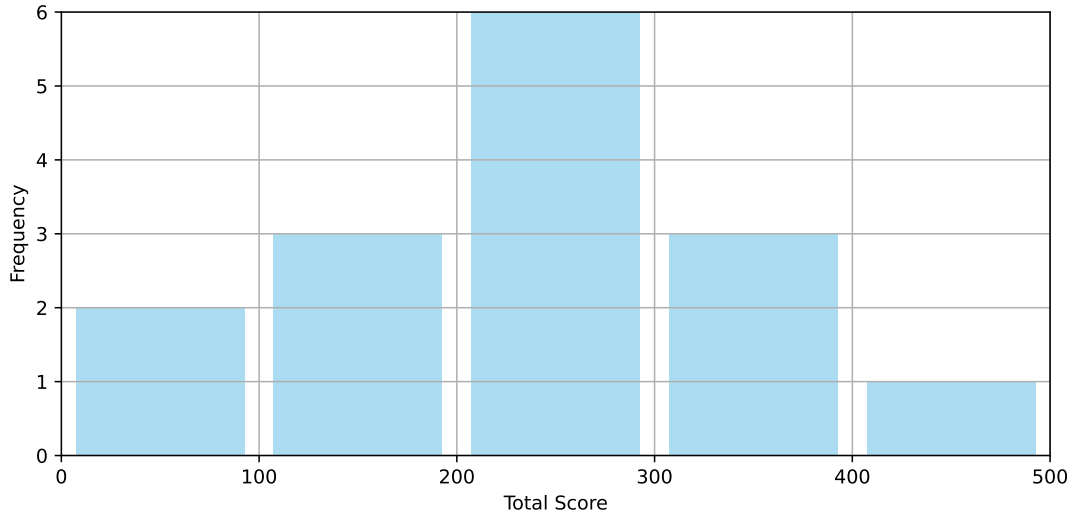


Figure 1.1.: Histogram of score distribution across tasks

As outlined in **Rule 4.5**, our approach to achieve a fair and even distribution of point scores, to accomodate teams at varying experience levels, was successful. Figure 1.1 shows a normal distribution of total scores achieved in ARCh 2026. However, Table 1.1 and the box plot shown by Figure 1.2 indicate how scores of 0 are still evident. This indicates that several teams are still encountering difficulties related to insufficient preparation in the leadup to ARCh, or critical issues preventing any task participation.

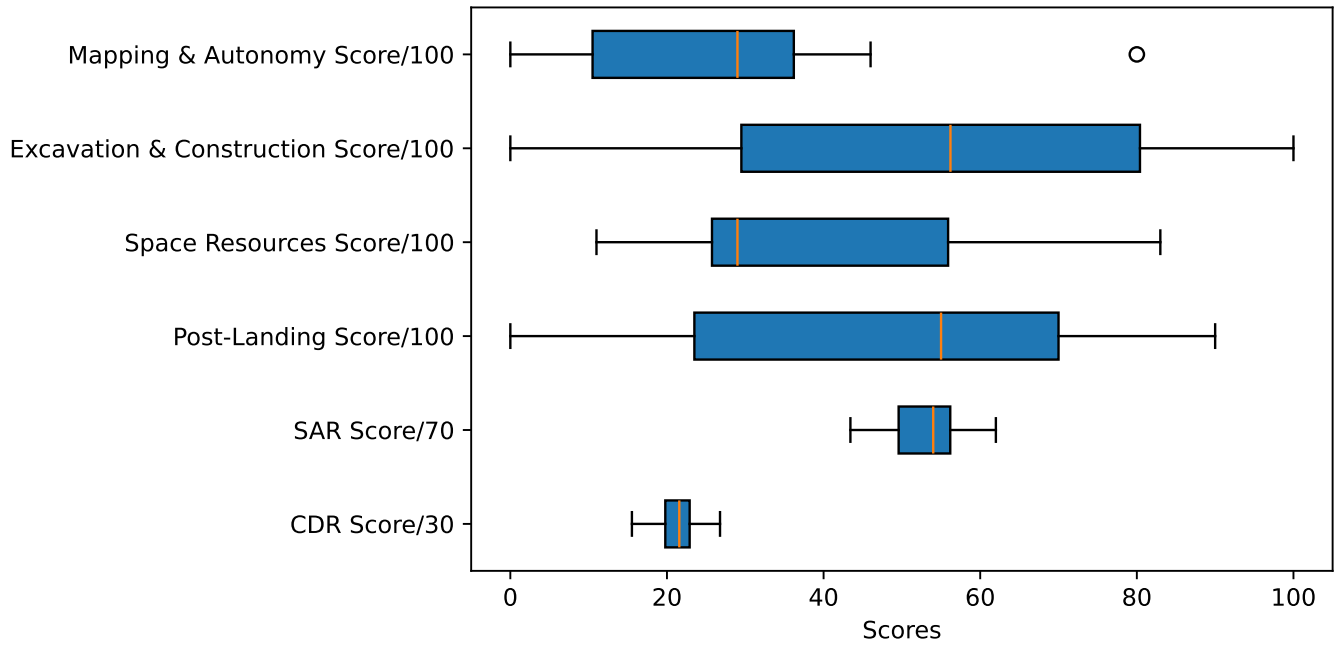


Figure 1.2.: Box and whisker plot of scores across tasks

Table 1.2.: Box and whisker plot statistics for task scores

Activity	Min	Lower Whisker	Q1 (25%)	Median (Q2)	Q3 (75%)	Upper Whisker	Max
CDR Score	15.53	15.53	19.79	21.57	22.90	26.78	26.78
SAR Score	43.43	43.43	49.58	54.00	56.18	62.00	62.00
Post-Landing Score	0.00	0.00	23.50	55.00	70.00	90.00	90.00
Space Resources Score	11.00	11.00	25.75	29.00	55.90	83.00	83.00
Excavation & Construction Score	0.00	0.00	29.50	56.20	80.40	100.00	100.00
Mapping & Autonomy Score	0.00	0.00	10.50	29.00	36.20	46.00	80.00

Table 1.3.: Team rover weights by task (kg)

Team	Post-Landing	Space Resources	Excavation & Construction	Mapping & Autonomous
UoA	46.00	54.40	45.00	38.6
RMIT	48.00	48.80	52.80	–
Monash	49.00	47.40	49.80	35.6
UniMelb	37.80	50.60	42.20	30.4
UTS	49.20	39.20	40.80	–
UQ	48.20	–	40.80	37.2
Scorpio	49.00	54.80	59.60	37.0
QUT	49.20	56.80	54.80	40.0
UNSW-RAS	45.80	50.20	51.00	40.0
UoW	41.60	32.00	47.80	27.0
AGH	45.00	56.60	58.20	45.6
Legendary	41.20	51.80	56.80	37.6
Bluesat	28.00	41.00	–	–
Deakin	49.00	29.20	–	–
UWA	–	50.40	–	–
Minimum	28.00	29.20	40.80	27.0
Average	44.79	47.37	49.97	36.9
Maximum	49.20	56.80	59.60	45.6

All teams conformed to the weight requirements described by **Rule 3.5**, with a 50 kg limit for Post-Landing and Mapping & Autonomy, and 60 kg for Space Resources and Excavation & Construction.

However it is evident that many teams are optimizing to fall at, or just under the weight limit for each task, rather than optimizing to *minimize* mass. For this reason, we have introduced an Efficiency Multiplier (see **Appendix D** in 2026 Rules & Regulations) which rewards teams that minimize size, mass and power. This was done with two clear objectives to i) reward teams that took steps to minimize mass below the task limit, and ii) reward teams which minimize size and power, in alignment with good industry practice.

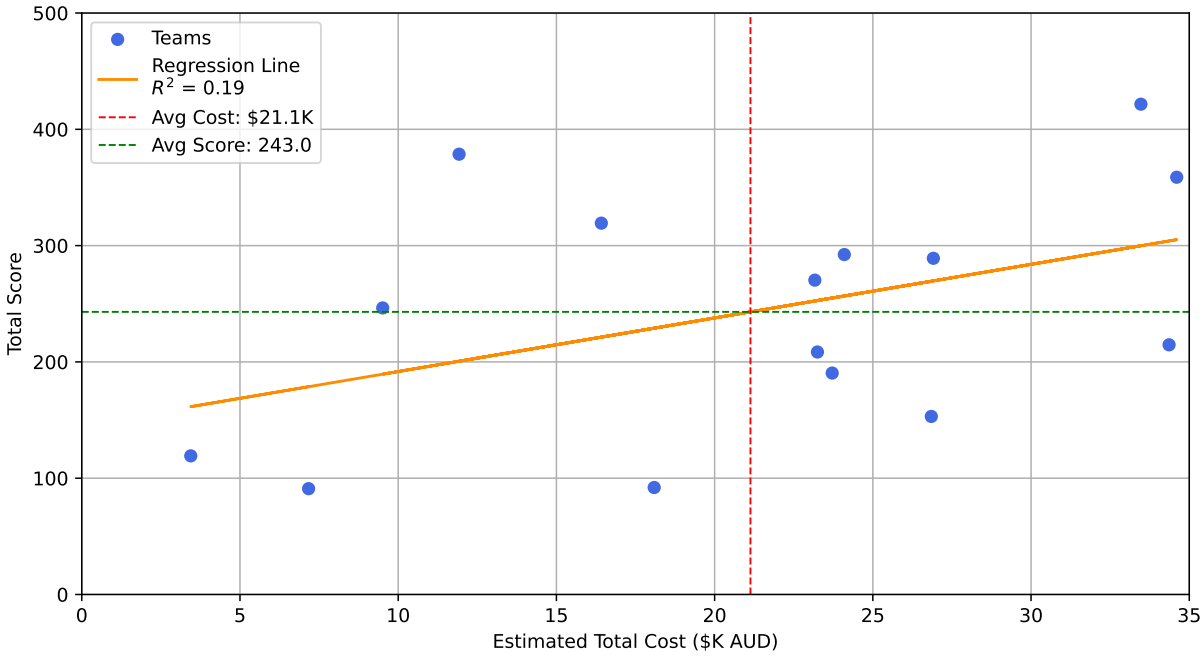


Figure 1.3.: Regression of total score with total cost ($K AUD)

Although Figure 1.3 indicates there is evidence of a positive correlation of total score and total cost of the rover, the R^2 is very low (0.19), and there is clear evidence of teams scoring above-average scores (> 243) with below-average rover costs (<AUD$21.1K). This clearly shows that increased funding does not necessitate improved performance, and that teams should more greatly consider the efficiency of how they utilize time, and money, towards outcomes.

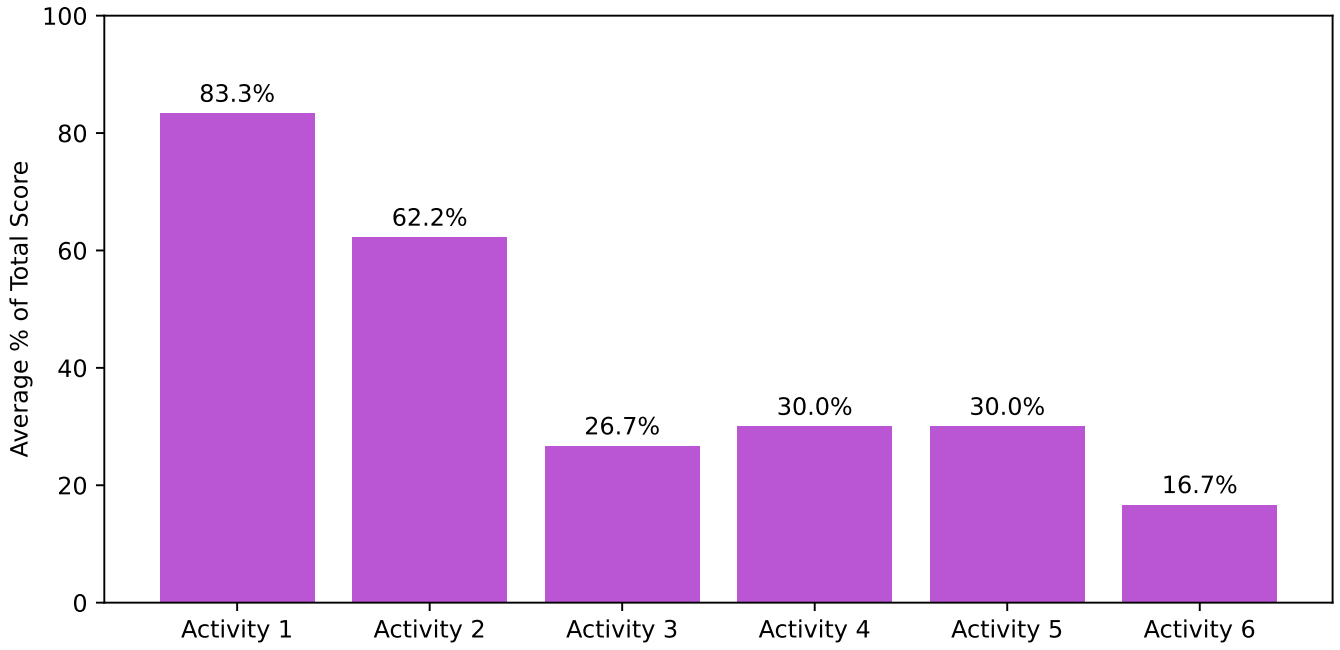
Part II.

Task Performance

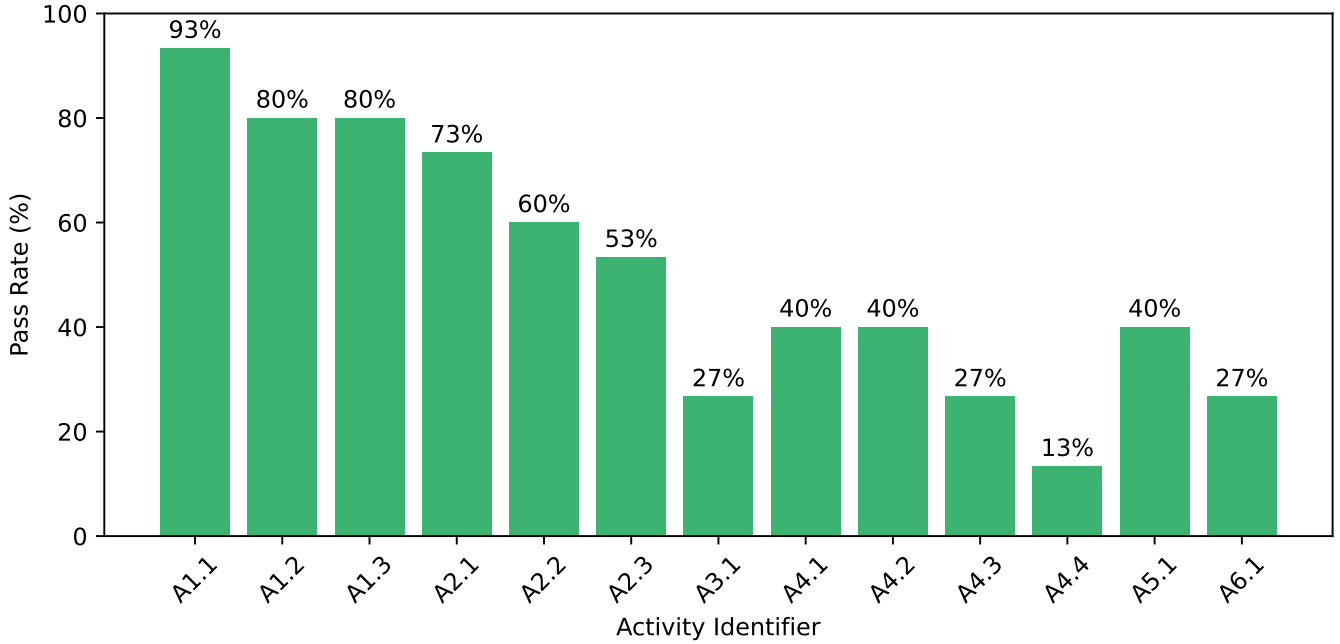
2. Post-Landing Task

There is strong participation with over 80% of teams leaving the start gate on Day 1 or 2, however it is expected that as teams mature and participate in their second ARCh this approaches 100%. For teams that do leave start gate, multiple points were collected for traversing to each gate. Difficulties arose for activity 3 and onwards which require specific capabilities of a robotic arm to interact with the processing plant and propellant hoses. As noted qualitatively by judges and evident in Figures 2.1a and 2.1b, if a team is able to leave the start gate on time and not waste time troubleshooting faulty systems during their task time, they acheived more points. However, participation rates decay to <40% for the final maintenance jobs for activity 3 (Processing Plant Maintenace) and Activities 4 and 5 (Hose Connections), with averge scores below 30%.

Overall, the current difficulty for this task and its activities is considered sufficiently challenging wherby teams that take steps towards optimizing time-allocation and ordering each activities during their task time, making systems more robust and less prone to failure (resulting in time loss), operator training and clear team communication for base-station and ground teams should result in teams achieving full points in these activities.



(a)



(b)

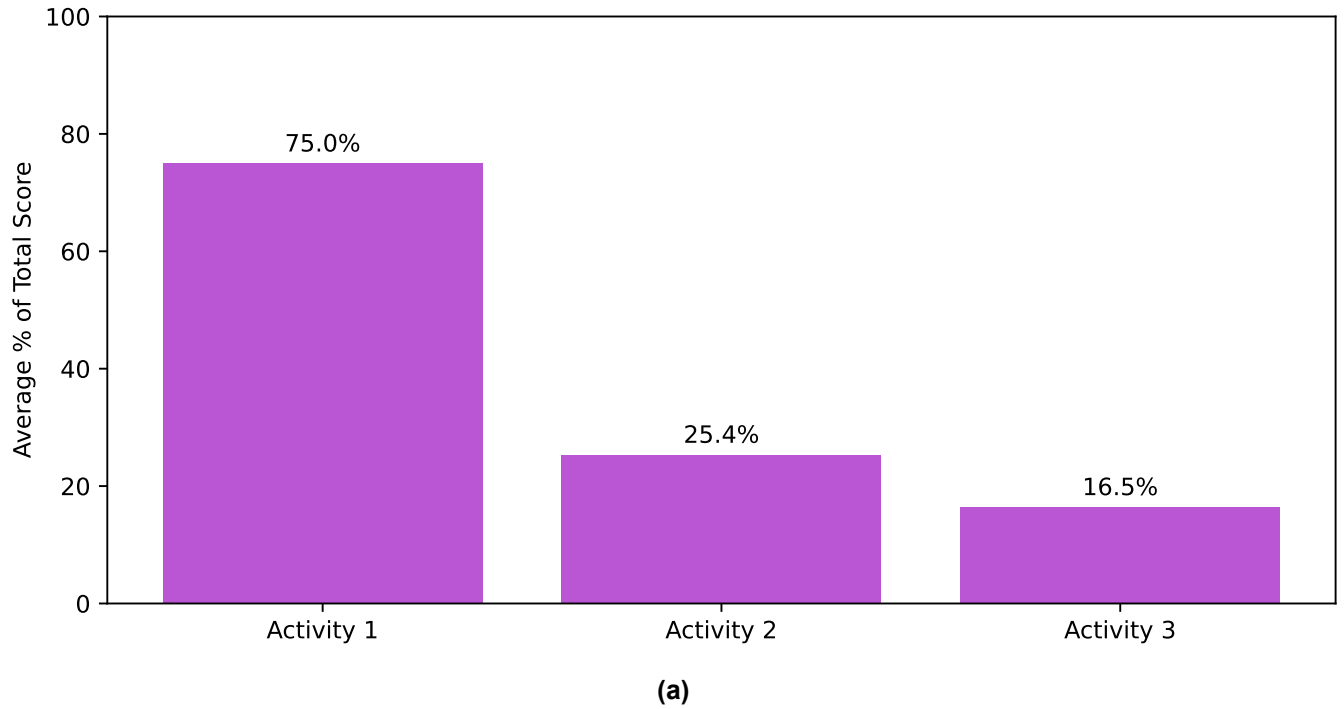
Figure 2.1.: Performance breakdown of Post-Landing (a) average score (% of Total Available) and (b) pass rate (%) for each activity.

3. Space Resources Task

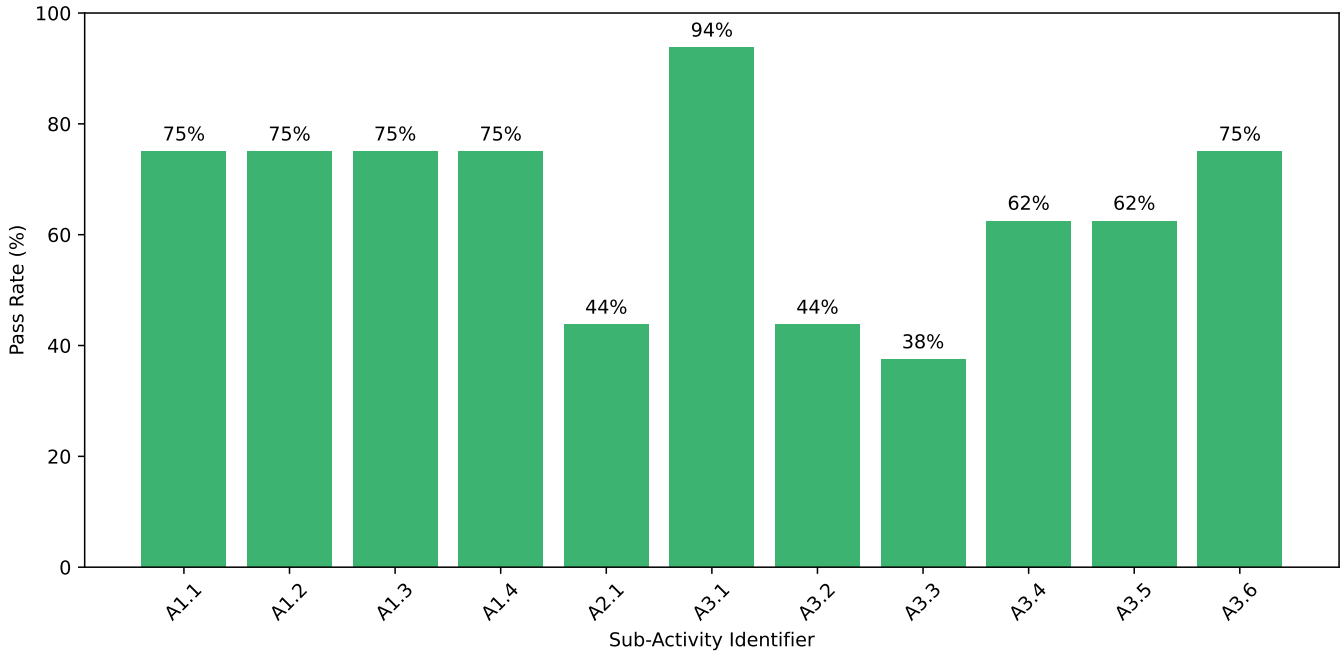
Overall, there was an increase in capability of mature teams in the number of teams processing water, but also the total amount of water yield. However, not all teams were capable of traversing to and imaging each site of interest as part of activity 1. This represents a growing dichotomy between new and existing teams, that is expected to be bridged as new teams mature. Estimates of ice and ilmenite content shown in Figure 3.2 teams are generally consistent with actual values. However, many teams continue to provide estimates that are outside the 0 to 30 wt.% mass of ice content and 0 to 15 wt.% mass limits that are **clealy** stated in **Appendices B.2**. This has occured for three consecutive years and indicates teams continue to not read the rules in their entirety.

For activity 2 (processing) there has been significant progress in the number of teams (7 teams) who successfully extracted water compared to 2024 (4 teams). Interestingly, the highest mass of water extracted in 2024 (9.07 g) was nearly double than the highest amount extracted in 2025 (4.86 g). Considering 2025 had sample sites with higher ice content (14.50 and 17.30 wt.%) compared to 2024 (13.65 and 8.35 wt.%), this suggests backward progression for some teams (i.e. less max water), but improvements for others (i.e. higher pass rate). Figures 3.3 and 3.4 indicate this disparity in team performance, with more than half of teams extracting no water at all, whilst two outliers yielded 2.49 and 4.86 g.

Overall, the current difficulty for this task and its activities is considered sufficiently challenging, only requiring that teams improve the efficiency of their processing and excavation sub-systems to maximize yield.



(a)



(b)

Figure 3.1.: Performance breakdown of Space Resources (a) average score (% of Total Available) and (b) pass rate (%) for each sub-activity.

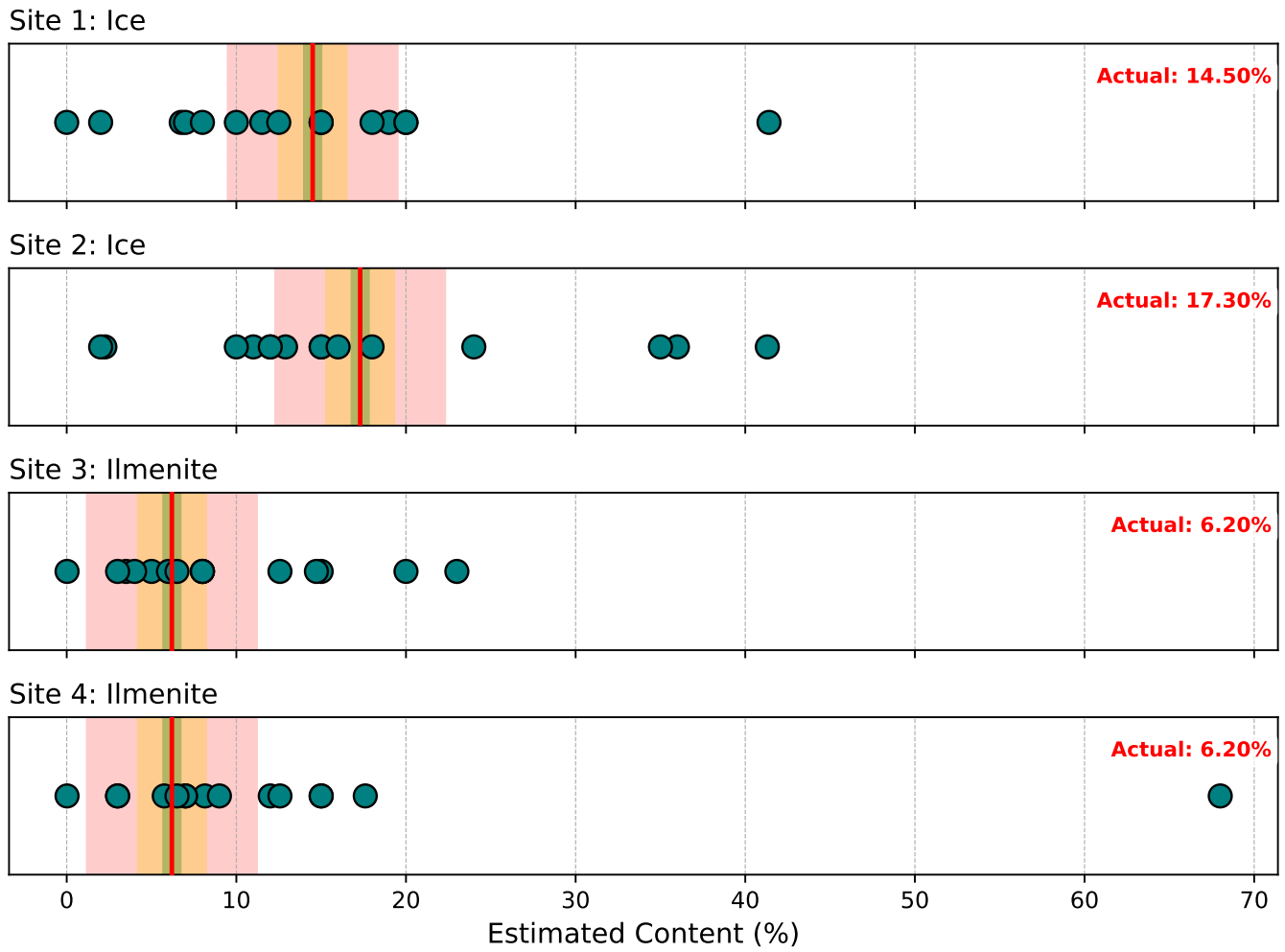


Figure 3.2.: Estimates of ice and ilmenite content (wt.%) for activity 1 (prospecting) of the Space Resources Task. Each dot represents a team's estimate for a given site, with the actual value shown as a solid red line and given in top right of each sub-plot. Shaded bands indicate accuracy zones relative to the true value: ±0.5% (green), ±2% (yellow), and ±5% (red) representing 5/5, 3/5 and 2/5 points scored respectively.

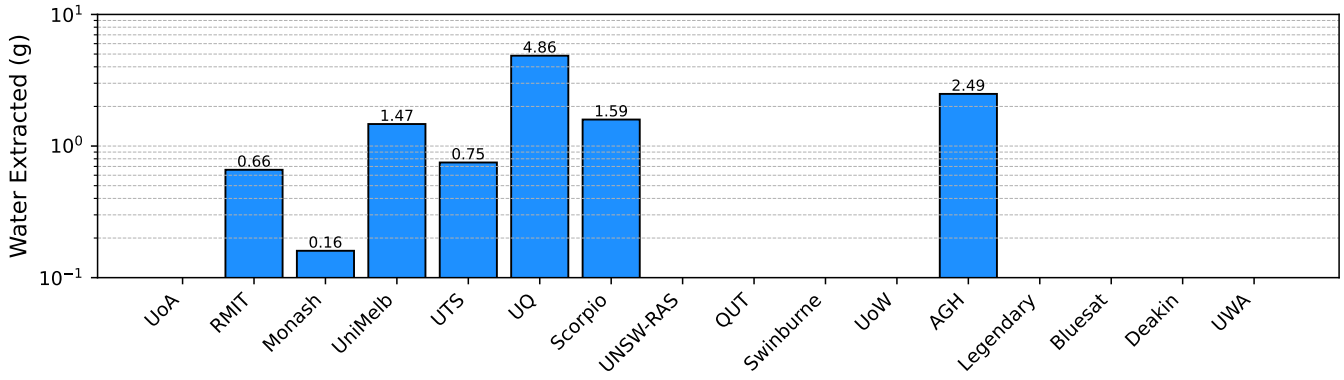


Figure 3.3.: Mass of extracted water for activity 2 (processing) of the Space Resources Task.

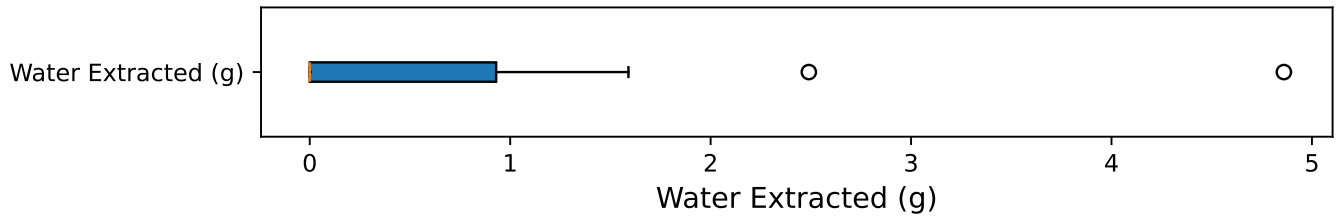


Figure 3.4.: Box and whisker plot of mass of extracted water for activity 2 (processing) of the Space Resources Task.

Table 3.1.: Box and whisker calculations for statistics for water extracted during Processing task

Metric	Min	Q1 (25%)	Lower Whisker	Median (Q2)	Q3 (75%)	Upper Whisker	Max
Water Extracted (g)	0.0	0.0	0.0	0.0	0.93	1.59	4.86

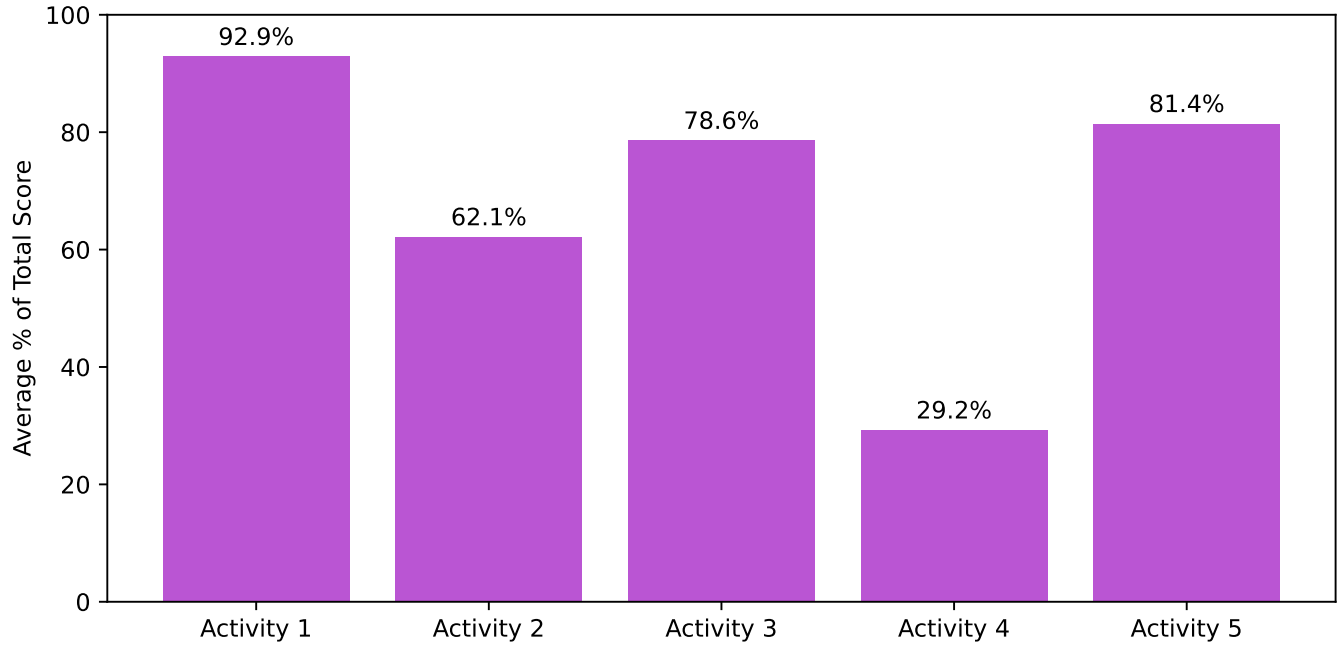
4. Excavation & Construction Task

There is strong participation with over 90% of teams leaving the start gate (activity 1) on Day 3 or 4, increasing from 80% for Post Landing task which occured on Day 1 and 2. This indicates progress made by teams during competition to fix issues related to basic rover functionality. Although welcome, judges urge teams to resolve these issues prior to arriving at the ARCh, as this will result in better outcomes at ARCh and also more time for field testing to identify issues related to rover functionality and team/operator training for smoother, stress-free operations.

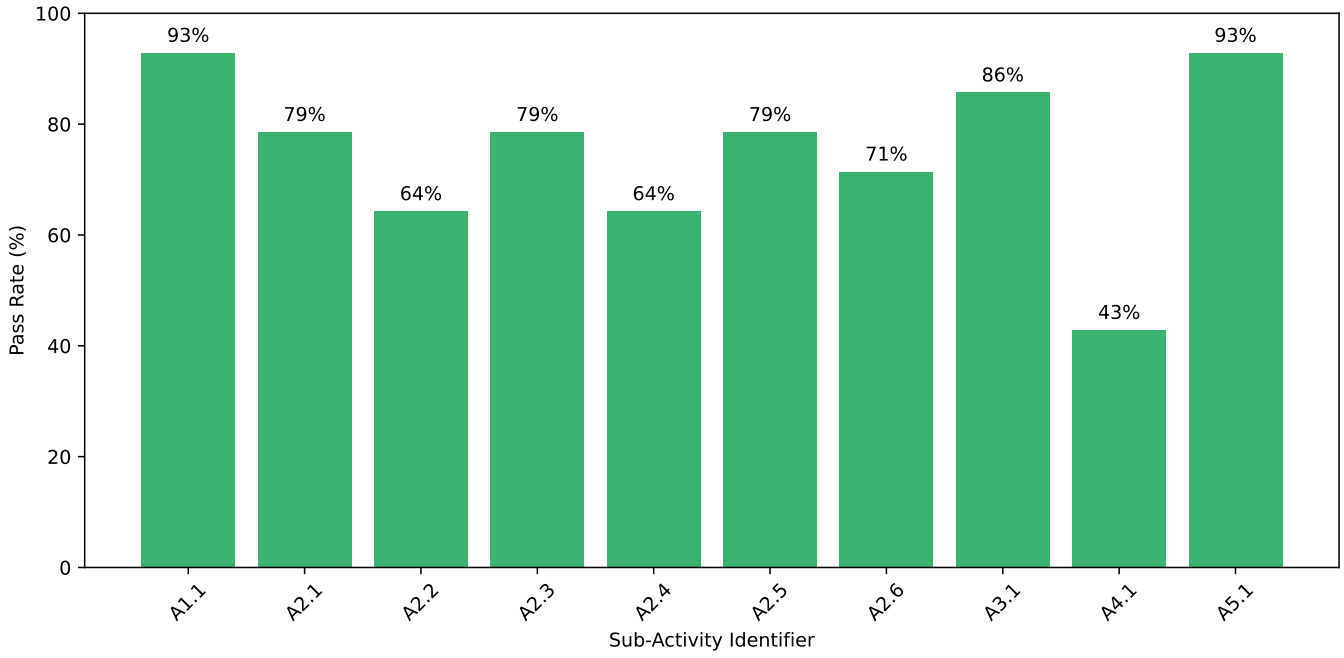
There was a very strong participation rate in activity 2 (rock clearing) with between 60 to 80% of teams successfuly clearing small, medium, large and huge rocks - with an average activity score of 62.1%. This success carried over to Activty 3 (berm construction) with 86% participation rate and average score of 78%. Although promising, as with activity 2 (processing) in the Space Resources task, there was a very large spread in the volume of regolith deposited, with an interquartile range of over 8483 cm^3 and more than half of the teams constructing less than 2608 cm^3—just 13.9 % of the maximum recorded volume of 18 808 cm^3. 11 of 15 teams (73%) constructed less than 9393.5 cm^3, suggesting the majority of teams fell below the 75th percentile for this task.

Activity 4 (paver construction) had the lowest pass rate (43%) with an average score of 29.2%, indicating it had the highest difficutly. This was largely due to most teams either not having a dedicated payload, and not attmepting this activity, teams running out of time and rovers getting stuck. No changes to this activity are required, and as teams mature it is expected the average score and pass rate will correspondingly increase.

As over 30 points are awarded for activity 2 (berm construction), representing a significant portion of total points for this task, the point allocation method has been modified. Rather than set bands as a function of volume, points in 2026 will be awarded as a threshold amount (5 points) with up to 25 points as a function of the largets quantity deposited by any team - in alignment with the method used for point allocation in activity 2 of Space Resources task- See **Rule 11.6.3.1** and **11.6.3.2**, and **Appendix C** for further details. This change is intended to avoid the need for judges to predict and modify volume estimates for point bands each year, makes scoring more competative, and should result in a better spread of scores.



(a)



(b)

Figure 4.1.: Performance breakdown of Excavation & Construction: (a) average score (% of Total Available), and (b) pass rate (%) for each sub-activity.

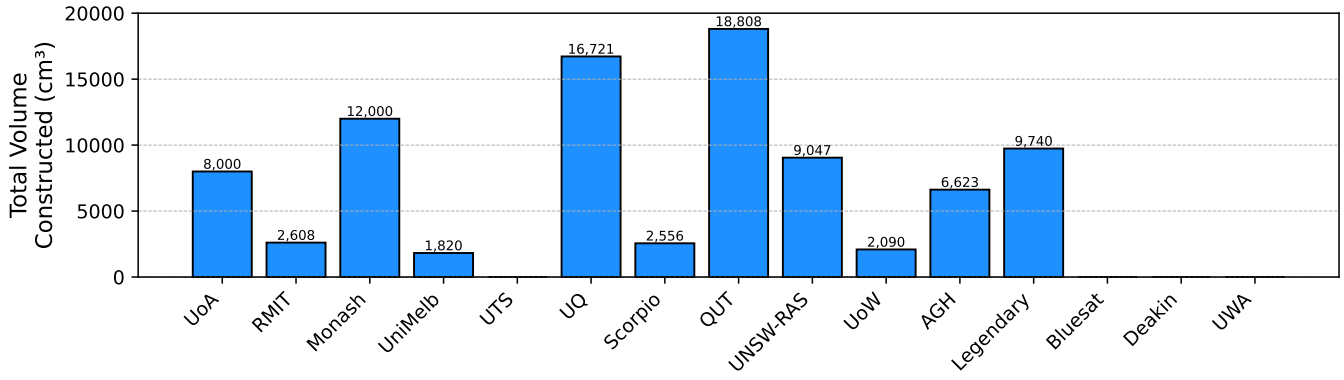


Figure 4.2.: Volume of berm constructed as part of activity 3 (Berm Construction) of the Excavation & Construction Task.

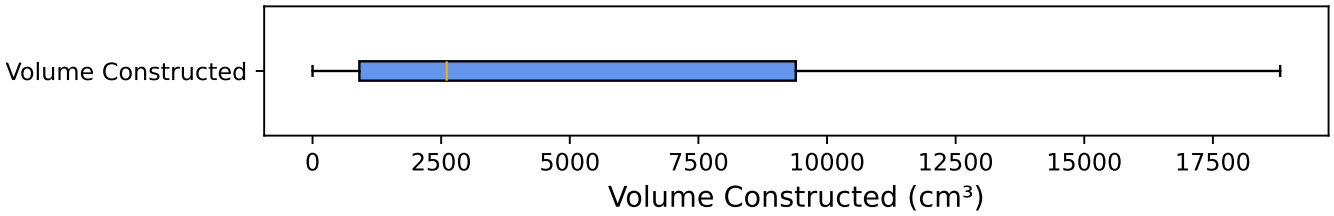


Figure 4.3.: Box and whisker plot of berm volume for activity 3 (berm construction) of the Excavation & Construction Task.

Table 4.1.: Box and whisker plot statistics for volume of regolith deposited during the Excavation & Construction task

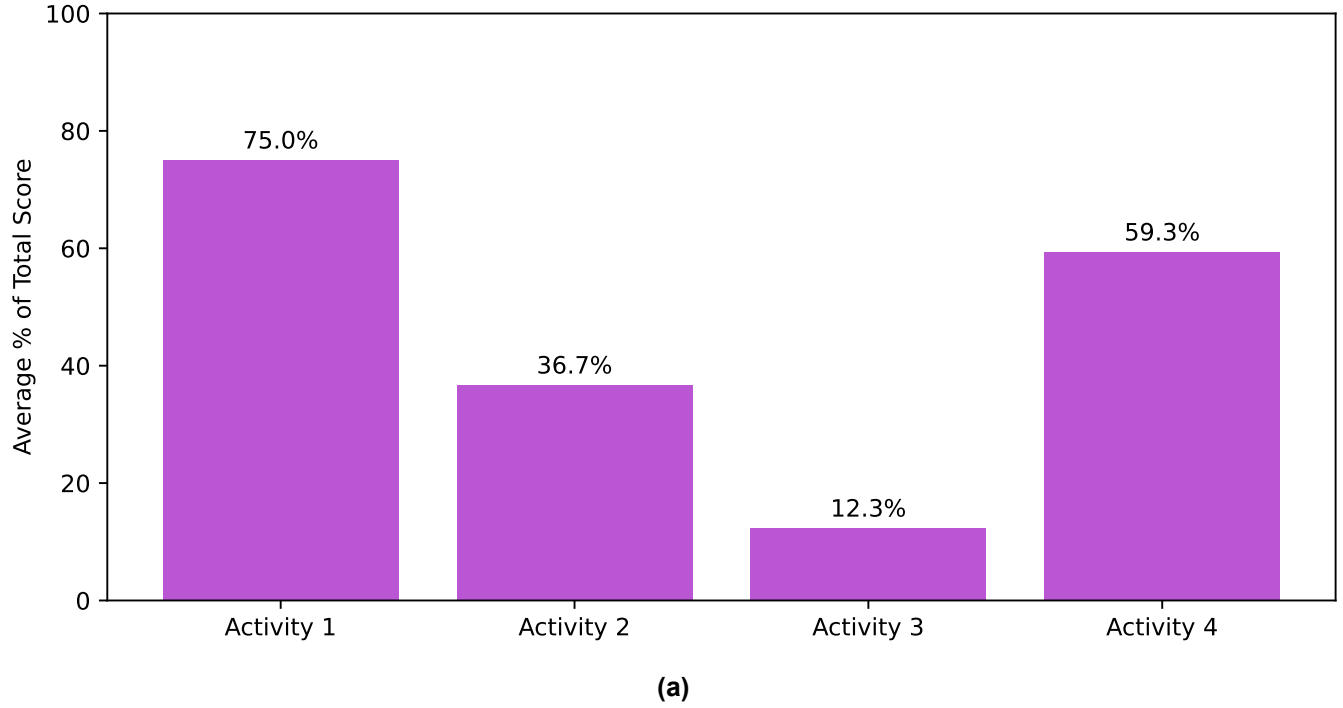
Metric	Min	Lower Whisker	Q1 (25%)	Median (Q2)	Q3 (75%)	Upper Whisker	Max
Volume Constructed (cm^3)	0	0	910.0	2608.0	9393.5	18 808	18 808

5. Mapping & Autonomy Task

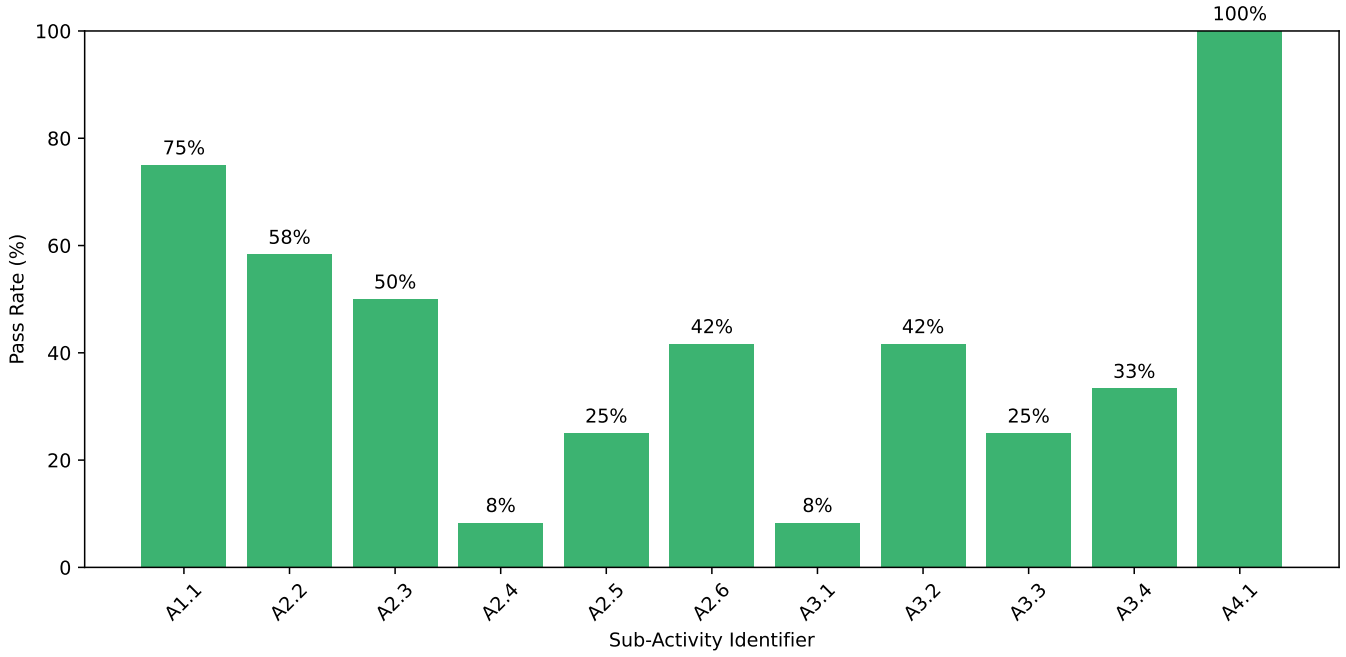
This task had the lowest pass rate, with only 75% of teams initiating the activity (i.e., leaving the start gate) with the primary sub-activities having pass rates below 50%, as shown in Figure 5.1b. As interpreted from the CDR and SAR deliverables handed to judges, autonomy continues to be a difficult sticking point, even for mature teams.

Activity 2 (landmark navigation) and activity 3 (exploratory mapping) had lowe average scores of 36.7% and 12.3%, respectively, indicating that most teams severely struggled with autonomous traversal and identification. To achieve any score for localize blocks in activity 3, errors had to be within 600 mm of the ground truth, with full points awarded only for localization within 300 mm. The IQR was broad, spanning up to 2.95 m for the white block for example, as illustrated in Figure 5.2, whilst the median error across all block colors exceeded the 600 mm threshold, resulting in zero points being awarded for at least half of the teams, whilst maximum errors exceeded 5 m, and in the case of the white block reached 14.21 m—equivalent to nearly 24 times the 600 mm margin required for any points.

Overall, this task continues to prove very challenging for teams however does not require modification by judges. Instead, teams should continue to improve their system and build on existing heritage and success each year using a near-identical platform, rather than undertaking a complete, or near-complete redesign and rebuild. As has occured multiple times in the past, this approach has resulted in poorer performance (or complete lack of participation) compared to the previous year, with added stress and difficulties of new systems with no prior demonstrated performance.



(a)



(b)

Figure 5.1.: Performance breakdown of Mapping & Autonomy Tasks: (a) average score (% of Total Available), and (b) pass rate (%) for each sub-activity.

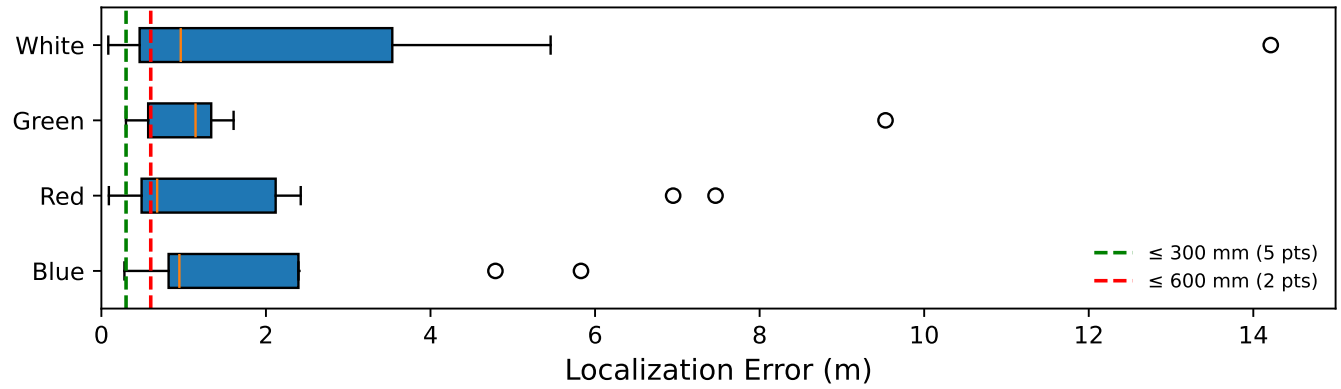
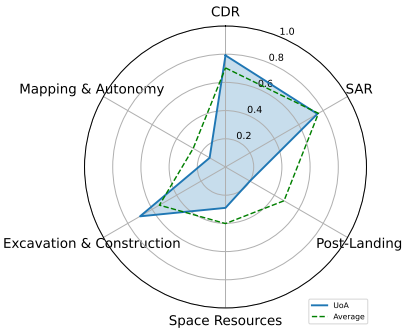


Figure 5.2.: Localization error for coloured blocks in activity 3 (exploratory mapping) of the Mapping & Autonomy Task.

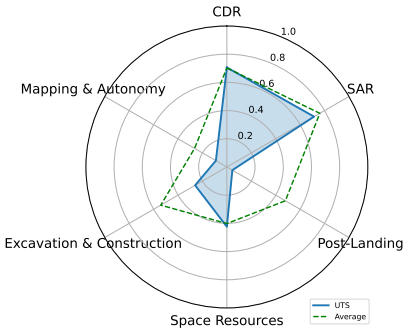
Table 5.1.: Box and whisker plot statistics for localization error during the Mapping & Autonomy task

Block	Min	Lower Whisker	Q1 (25%)	Median (Q2)	Q3 (75%)	Upper Whisker	Max
Blue	0.280	0.280	0.817 00	0.9480	2.395 00	2.395	5.829
Red	0.091	0.091	0.488 75	0.6780	2.118 25	2.423	7.464
Green	0.299	0.299	0.568 50	1.1445	1.335 00	1.607	9.530
White	0.084	0.084	0.464 75	0.9645	3.535 75	5.460	14.212

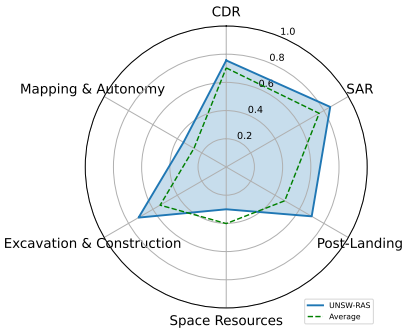
Appendices



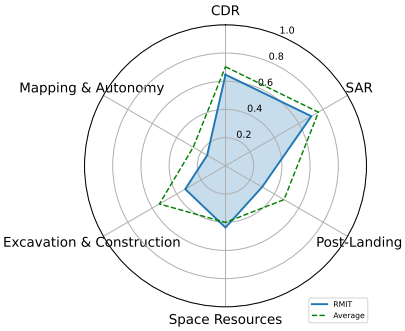
UoA



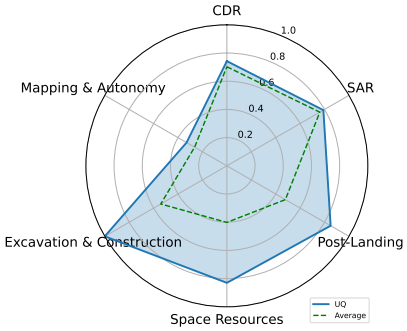
UTS



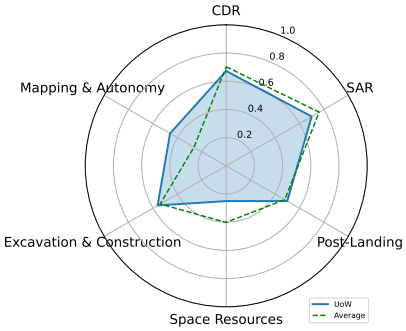
UNSW-RAS



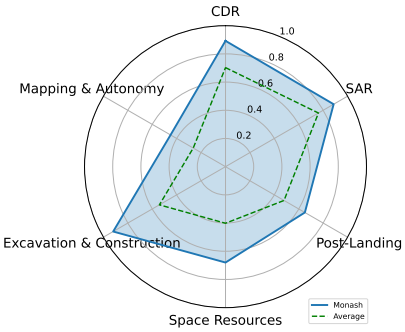
RMIT



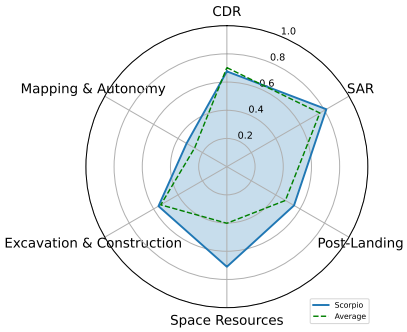
UQ



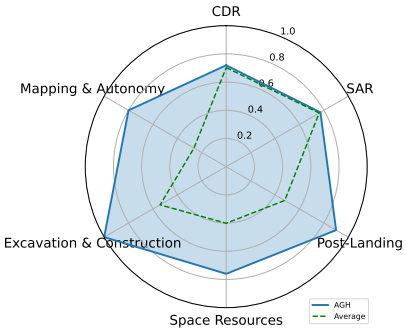
UoW



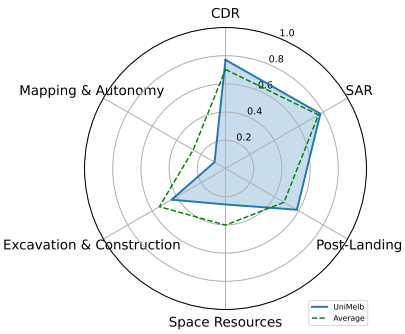
Monash



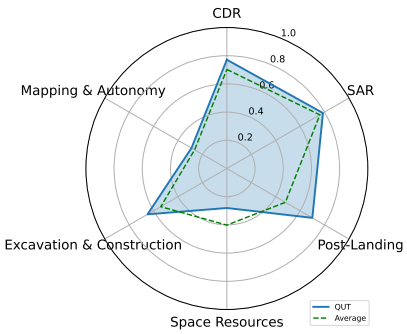
Scorpio



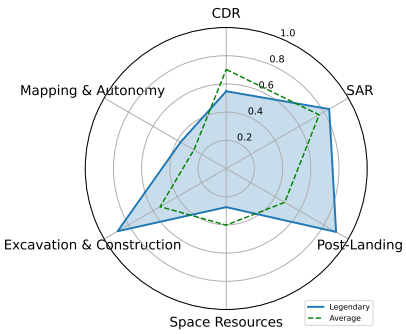
AGH



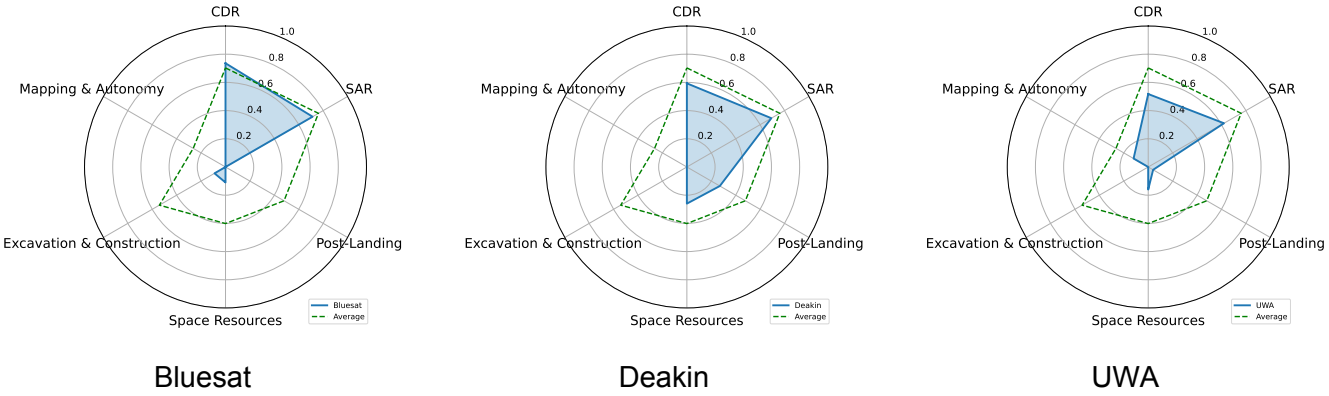
UniMelb



QUT



Legendary



Bluesat Deakin UWA

Figure 3.: Radar charts showing normalized scores for each team across all competition categories. Green overlay indicates average performance across all teams.

??