

10 Years of IREC: A Quantitative Scoring Analysis

Carl Anderson

October 16th, 2026

Purpose

The purpose of this document is to present a quantitative analysis of the published scoring results from ten Intercollegiate Rocket Engineering Competitions (IREC) between 2016 and 2026. IREC has been run as the Spaceport America Cup since 2017, and 2016 was not the first competition (the organizers list the 2019 event as the 14th), so these ten are the most recent part of a longer series and not its start. The results from each year are reconciled into a single dataset, which is then used to examine how overall and component scores, competition classes, individual schools, and category winners changed over that period, and to state which of the observed trends the data support and which they do not. The 2020 competition was cancelled and 2021 was held virtually, so there are ten sets of results rather than eleven.

Executive Summary

The results for each competition were published as a separate file (nine Excel workbooks and one PDF), and no two of them share the same layout. The scoring rubric, the class names, and the way missing data was recorded all changed from year to year. After reconciling them there are 1,031 team entries from 289 schools in 30 countries, and every row that has a component breakdown (935 of 935) adds up to its reported total to within 0.03 points. Because this analysis is based only on what the organizers published, its quality is limited by those files. They list scored entries rather than registrations, the rubric changed several times, and 2018 and 2021 have less detail than the other years. More can be read on the steps behind the analysis in the Methodology section. Despite these issues, I believe the files support a useful look at long term trends, as long as results that are close to the noise are treated that way.

Beginning with the field, entries grew from 82 in 2017 to 143 in 2025 and 140 in 2026 (fig. 1), mostly in COTS. The median base score (total points excluding bonuses, as a share of that year's maximum) went from 44% in 2017 to 72% in 2026, a slope of about +2.2 points per year over the comparable years (95% CI +1.7 to +2.8, fig. 3). The increase is in the middle and bottom of the field, since the 90th percentile and the winning score have not moved since 2022. Two changes seem to account for most of it: more entries recorded a valid apogee (56% in 2018 and 83% in 2026), and paperwork and design scores rose and became less spread out. As a result, flight results now explain most of the differences between teams' total scores, and the gap between first and tenth place fell from 22% of the winner's score in 2019 to 4.8% in 2026 (fig. 8). Part of that is probably mechanical, since paperwork scores are close to their caps, but it does mean small differences in flight performance now decide most of the placings.

The competition classes did not move together. Comparing 2017 and 2019 against 2022–26, the median base score rose by 21 to 31 points in 10k COTS, 10k SRAD solid, and 30k SRAD solid, but

by only 2 to 4 points in 30k COTS and SRAD hybrid/liquid (fig. 10), and the 30k COTS change can't be distinguished from zero. As a result, 10k COTS now leads 30k COTS by a median of 21 points, compared to 3 points before the pandemic gap. The flight record points the same way: in 2022–26, 83% of 10k COTS entries recorded a valid apogee compared to 61% of 30k COTS entries. No class is detectably over or under represented in the top ten places (permutation test, $p = 0.58$), although with only 50 places that test can only pick up large differences.

Two further questions came out of the component results. First, design and build scores are a weak predictor of flight results, and a better one than the technical report scores (fig. 7). A team with a valid flight outscored a team without one on design and build 60% of the time, compared to 55% for the report (50% would mean no information), but among the teams that flew neither score is related to how well they flew. The link was clearer in 2019 and 2022 than in 2023–26, when design and build scores were bunched closer to the cap. Second, the share of entries earning the 50 point CubeSat bonus went from 38% in 2019 to between 56% and 63% in every competition since 2023, and entries with the bonus stand on average 14 percentile points higher on the base score than those without it (fig. 20). The data can't say whether the payload helped or whether stronger teams were more likely to build one.

At the school level the results are noisy. A school's standing at one competition has a rank correlation of only $\rho = 0.28$ with its standing at its next appearance (paperwork and design 0.46, flight score 0.28), and schools in the top fifth of the field average the 65th percentile the next time they appear (fig. 13). The estimated gain per additional appearance is +1.6 percentile points (95% CI -0.6 to $+3.8$), which can't be told apart from zero. Class changes are common: 28% of consecutive appearances are in a different class, and about two thirds of those are a step up in altitude or from COTS to SRAD. Schools that did not record a valid apogee were more likely to move from SRAD to COTS the next year (20% against 6%) and less likely to move up from 10k COTS to 30k or higher (7% against 21%). Each of these comparisons is borderline ($p = 0.028$ and 0.047), and once the four class-switching tests are adjusted for each other none is below 0.05, so they should be treated as suggestive (fig. 16).

There have been nine declared overall winners from 2017 to 2026, from eight schools, and 2016 had separate Basic and Advanced winners instead. The University of Sydney won in 2022 and 2025 and every other winner won once, so no school has won at consecutive competitions. The overall winner is chosen by the organizers from among the category winners using qualitative as well as quantitative judgement, and it was the highest-scoring entry in five of the nine years (not in 2017, 2019, 2023, or 2025). Category titles repeat more often: the same school won a category at consecutive competitions in 10 of 50 cases, against about 3 expected if each year's winner were drawn at random from the entries (simulation, $p < 0.001$, fig. 22). All 25 winners of categories with at least 5 entries since 2022 recorded a valid apogee and recovered, and their median altitude miss was 1.9% compared to 9.5% for the rest of the field.

The results above are not all equally certain. The main reasons are:

- **Rubric changes.** Scores are put on a common percentage scale, but that does not account for differences in what was being judged or how strictly. The paperwork and design trend in particular may be partly judging calibration instead of team quality.
- **Small classes.** SRAD hybrid/liquid, 30k SRAD solid, and multi-stage have between 5 and 17 entries per year, so their year by year values have wide intervals. For those classes the pooled before and after comparison is more reliable than any single year.
- **Correlation, not causation.** The class switching, design and build, and CubeSat results describe how entries relate to each other. Weather, rule changes, and team composition are

not in the data.

- **Multiple tests.** The four class-switching comparisons were each tested separately. After a Holm correction for running four tests, none has an adjusted p below 0.05 (they range from 0.11 to 0.41), so they are suggestive only.

Analysis

This section goes through the results in roughly the order the questions come up: who competes and whether schools come back (fig. 1 and fig. 2), how overall and component scores moved (fig. 3 to fig. 6), whether paperwork scores predict flights (fig. 7), how close the top of the field is (fig. 8), how the classes compare (fig. 9 and fig. 10), how individual schools do over time (fig. 11 to fig. 13), when schools change class (fig. 14 to fig. 17), how accurate the flights are (fig. 18 and fig. 19), the CubeSat bonus (fig. 20), and who wins (fig. 21 to fig. 24 and the winner tables). Unless noted otherwise, *base score* is the total score excluding bonuses as a share of that year’s base maximum, and *standing* is a school’s percentile within that year’s field (1 is the winner and 0 is last). Hollow markers in the line charts are years that aren’t directly comparable (2016 used a different rubric, 2018 reports totals only, and 2021 was virtual), and the grey band marks the cancelled 2020 competition. Definitions are in the Methodology section.

The field

Figure 1 shows the number of entries at each competition, split by propulsion family in (a) and by region in (b). The 2016 competition used Basic and Advanced classes instead of altitude and motor classes, so it appears as one grey block in (a). The entries here are rows in the scoring files, which can include teams that did not launch, so they are best read as scored entries rather than registrations.

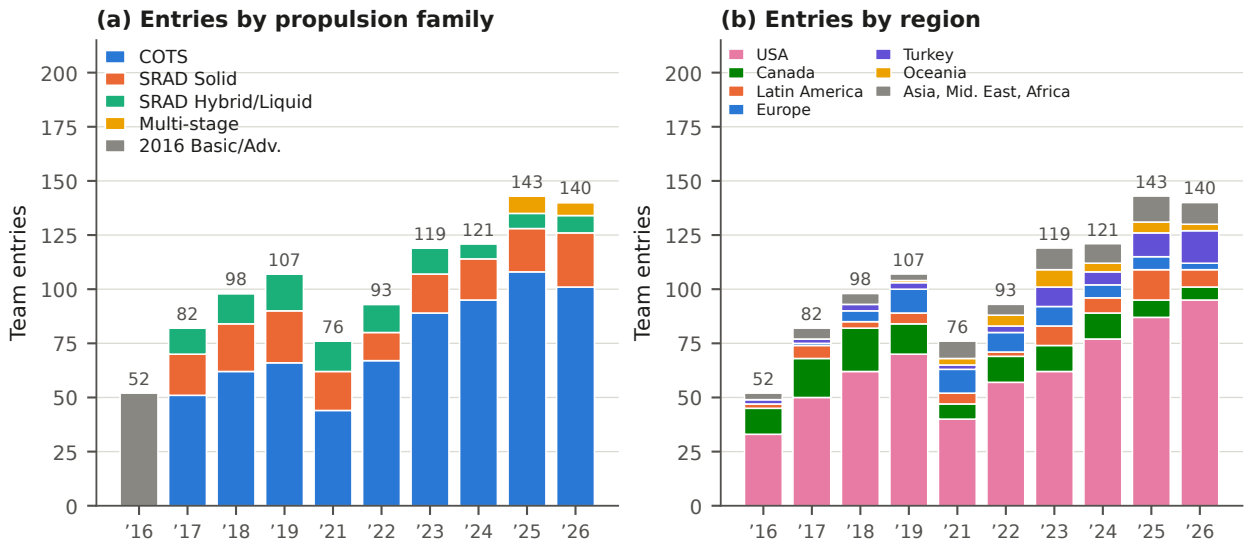


Figure 1: Entries per competition by propulsion family (a) and by region (b).

Entries went from 82 in 2017 to 143 in 2025 and 140 in 2026. Most of the net increase is in COTS (51 entries in 2017 and 101 in 2026), and most of that is in 10k COTS (45 entries in 2017

and 80 in 2026, with a peak of 87). SRAD solid entries were roughly flat (19 in 2017 and 25 in 2026) and SRAD hybrid/liquid entries went from 12 to 8. Between a quarter and a third of entries have targeted 30k or 45k in every competition since 2018.

Some schools fielded more than one team: 4 schools in 2016, 12 in 2017, and 11 in 2018, none in 2019, 2021, or 2022, and one or two per year since 2023. This follows the organizers' rules on teams per school. A limit on entries per student organization was removed in May 2018 after participant lobbying, the April 2022 draft rules limit competing teams to one per university because of demand for team slots, and in October 2024 ESRA announced that two teams from one university would be allowed again for 2025 if they use different propellant types (rule 3.3.5 of the 2027 rules, see Organizer sources). I could not find which revision introduced the one team limit, and the files show two teams from Istanbul Technical University in 2023 and 2024, which I can't explain from the rules I have seen.

In (b), the regions are the US, Canada, and Turkey (the three largest national contingents), and the other countries grouped as Latin America, Europe, Oceania, and Asia, the Middle East and Africa (the countries in each are listed in the Methodology section). The US share of entries has stayed at about 60% (63% in 2016–19 and 61% in 2022–26), so US entries have grown with the field and not faster than it. Canada's share was 22% in 2017 and 20% in 2018 and has fallen to 4% in 2026. This most likely reflects the growth of Launch Canada, a Canadian competition that is more convenient and cheaper for Canadian teams to attend than the Spaceport America Cup, but the data can't test that. Turkey went from 2 entries in 2016 to 15 in 2026 and was the largest national contingent after the US in 2025 and 2026. Each of the other regions grew by a factor of 1.6 to 2.3 per competition between 2016–19 and 2022–26, and Oceania went from a single entry before the pandemic to about 5 per competition. The share of entries from outside the US has stayed between 32% and 48%.

Figure 2 shows how often schools come back. Outside the pandemic transitions, between 62% and 80% of one competition's schools appear at the next. The rate was lower after the pandemic gap: pooled over the pairs of competitions, 78% of schools came back in 2016–19 (95% CI 72% to 83%) and 69% in 2022–26 (65% to 73%), a difference of 9 points (Fisher exact $p = 0.019$). In 2022–25 schools in their first appearance came back less often than schools with an earlier appearance in the data (59% against 72%), but that doesn't account for the whole fall, since first appearances were a smaller share of each competition's schools than in 2018 (17% to 27% against 35%) and the 2018 figure is itself overstated because the data start in 2016. Only 48% of 2019's schools took part in the virtual 2021 competition, and 50% were back in 2022.

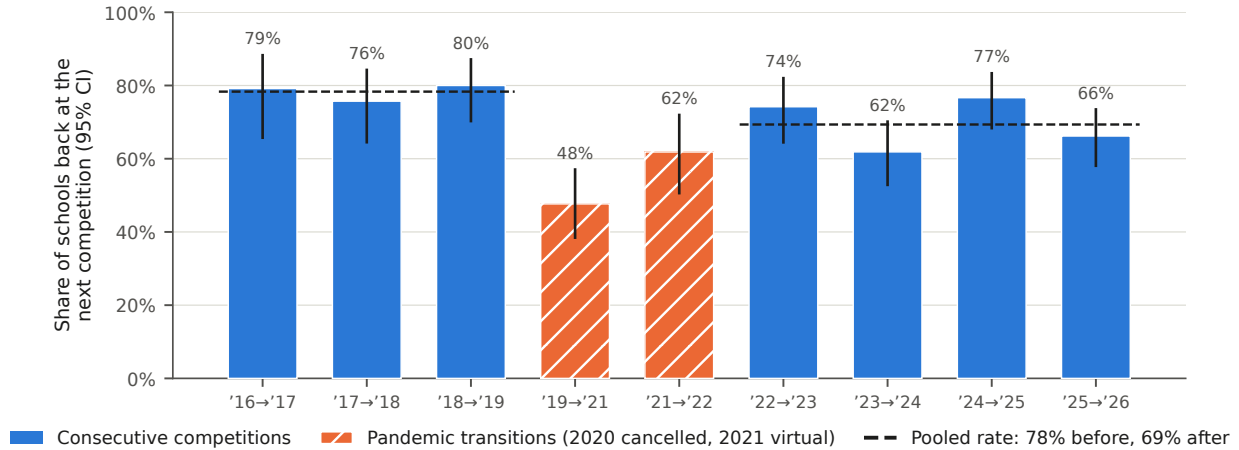


Figure 2: Share of one competition's schools that appear at the next, with the pooled rate before and after the pandemic gap.

Overall and component scores

Figure 3 shows the distribution of base scores at each competition. The line is the median entry with the middle 50% of entries as a band, and the other two series are the 90th percentile and the winner.

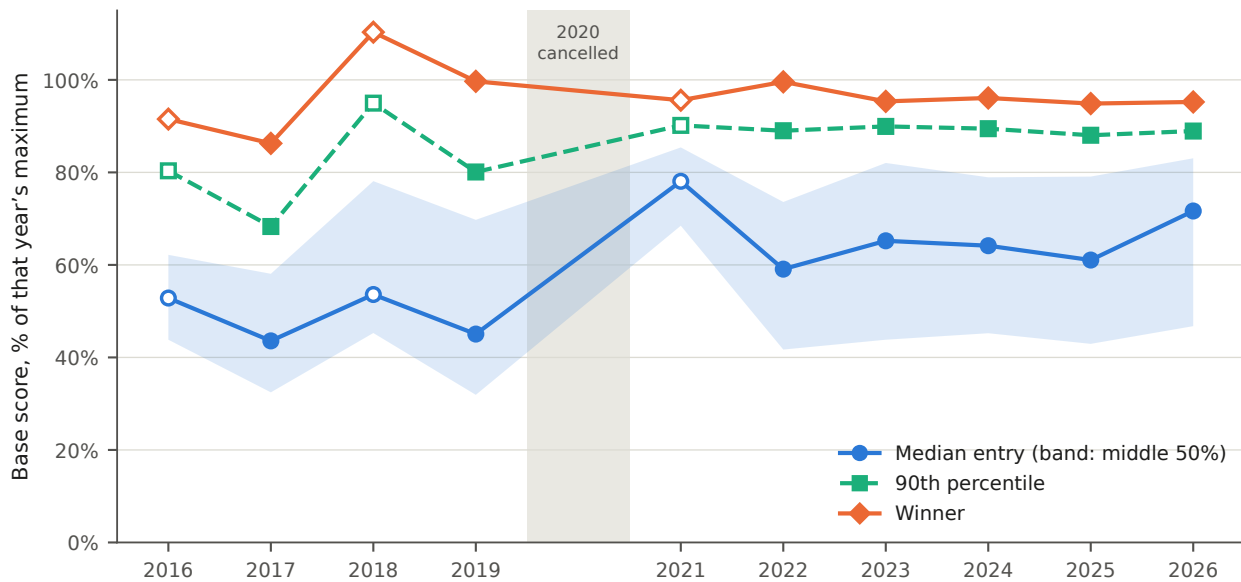


Figure 3: Base score as a share of the maximum: median, middle 50% of entries, 90th percentile, and winner.

The median went from 44% in 2017 to 45% in 2019, 59% in 2022, and 72% in 2026. Over the comparable years (2017, 2019, and 2022–26) the slope is +2.2 points per year (95% CI +1.7 to +2.8). The 90th percentile (88% to 90%) and the winner (95% to 100%) have been flat since 2022, so the increase is in the middle and bottom of the field. It should be noted that the 2018 values are overstated. That file only reports totals, so the bonus points can't be removed, which is why

the winner's value is above 100%.

Figure 4 breaks the score into its components. Admin points are at or near the cap: the median entry earns full marks in every year since 2017 (2018 has no breakdown), and between 69% and 96% of entries earn the full 60 points in 2022–26. The low end of that range is 2025, when 25 entries scored 45 (one 15 point item missing) and 13 received partial credit on one item, which doesn't happen in any other year. The median technical report score has been steady at 80% to 86% since 2022, although the weakest reports improved (bottom quartile 61% in 2019 and 71% in 2026). Design and build is the clearest riser, with a median of 74% in 2017 and 88% in 2026, and the middle half of entries narrowed from 31 points wide in 2017 to 12 in 2026. Flight scores among the entries that flew moved the same way, with a median of 51% in 2019 and 70% in 2026 and a middle half that narrowed from 65 points wide to 41. The all-entries median in (d) is lower because of teams with no valid flight.

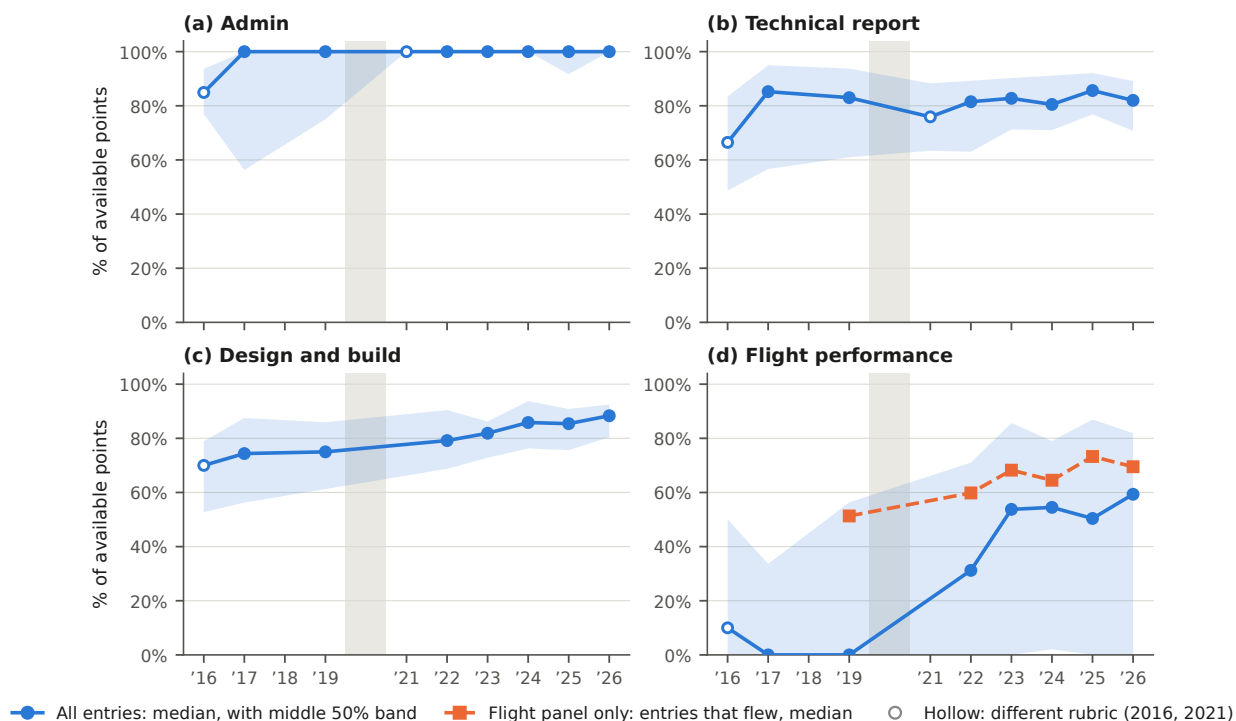


Figure 4: Median score and middle 50% band for each scoring component, as a share of available points.

Figure 5 shows (a) the share of entries that recorded a valid apogee and, of those, the share that were awarded recovery points, and (b) the standard deviation of the paperwork and design score and of the flight score across entries. The share with a valid apogee went from 56% in 2018 to 83% in 2026, a slope of +3.3 points per year (95% CI +2.3 to +4.4), and recovery among teams that flew went from 52% to 71%. The paperwork and design score also rose, from a median of 78% in 2019 to 86% in 2026, with the bottom quartile improving the most (57% to 80%). In (b) the spread of paperwork and design scores fell from 26% in 2019 to 10% in 2026, while the spread of flight scores across all entries did not (36% in 2026), since teams with no valid flight score zero. Judging and rubric details changed over the period (for example, 2023 split design quality and build quality), so part of the paperwork gain may be judging calibration, and I can't separate the two with this data.

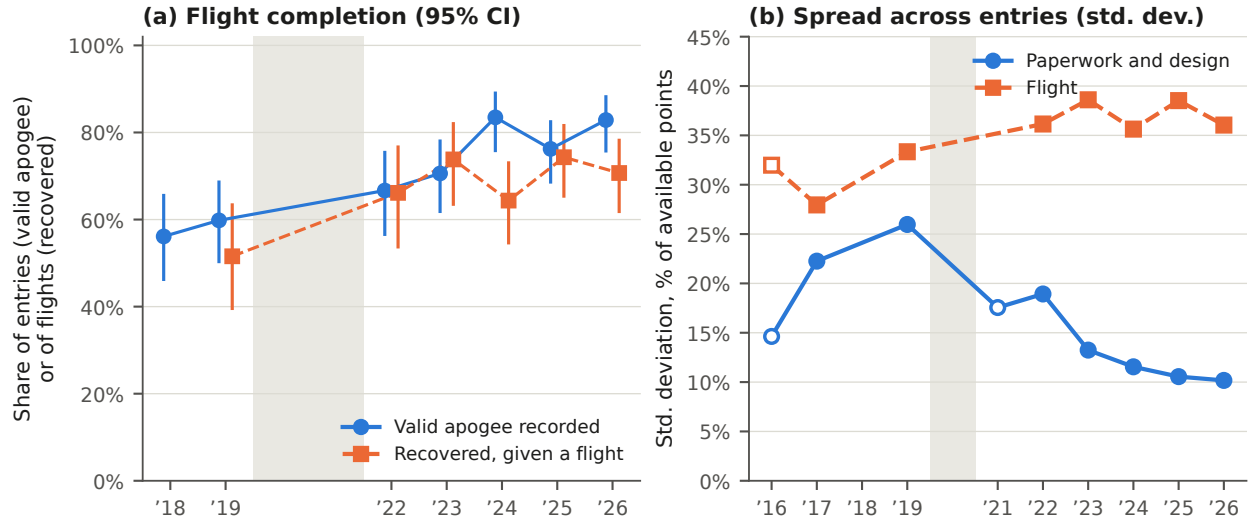


Figure 5: Flight completion and recovery (a), and spread of paperwork and flight scores (b).

Figure 6 shows how much of the difference between teams' total scores each component accounts for, calculated as the covariance of the component with the total divided by the variance of the total. Flight accounts for 65% to 79% of the variance in 2022–26, compared to 61% in 2017 and 51% in 2019, while report and design combined went from 35% of the variance in 2017 to 10% in 2026. Bonus minus penalty has been the second largest source of variance in every competition since 2023, at 18% to 25% in 2023–25 and 11% in 2026 (still ahead of design and build at 7%). It combines bonuses (prompt launch, sportsman, CubeSat) with penalties, which include late submission penalties of up to 345 points in 2024–25.

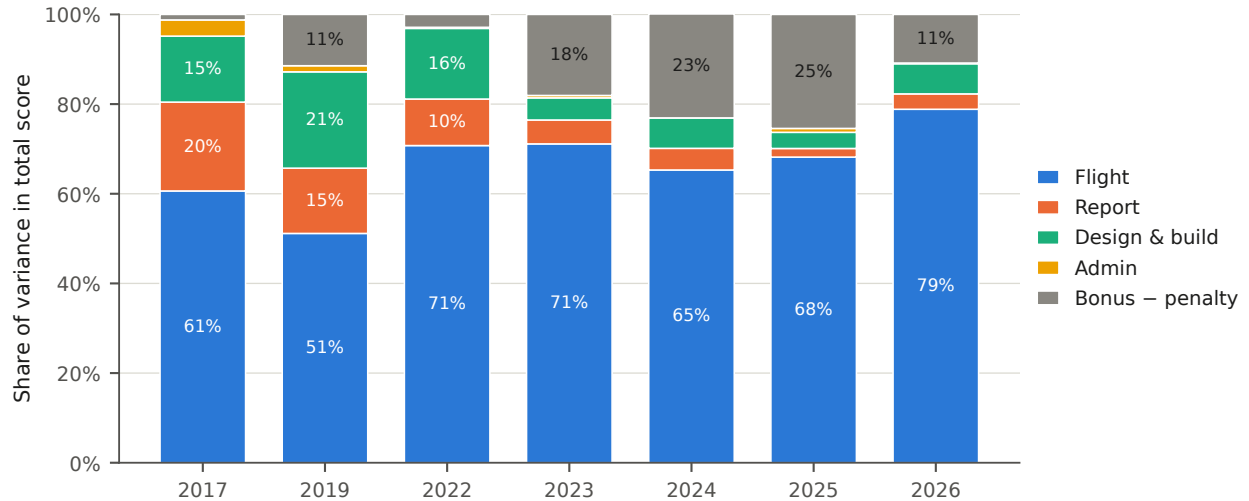


Figure 6: Share of the variance in total score accounted for by each component.

Paperwork scores and flight results

The component figures show design and build scores rising and bunching together over the period, and flight scores among the entries that flew doing the same. That raises the question of whether the scores are related for individual entries, and in particular whether design and build scores say more about how a team will fly than the technical report scores do. Each score is converted to a percentile rank within its competition so that rubric changes and score levels do not matter. The comparison covers 2019 and 2022–26, the years with both a component breakdown and flight outcomes (723 entries from 266 schools), since 2018 has no breakdown and 2021 had no flights.

Table 1 gives the results. Design and build has a small positive rank correlation with the flight score (0.11), and the technical report a smaller one that can't be told apart from zero (0.05). The difference between them is itself uncertain (0.11, 95% CI -0.03 to 0.24), so the data suggest that design and build is the better predictor of the two but do not establish it. Most of the association is about whether a team gets a valid flight at all. The bottom quarter of design and build scores had a valid apogee 62% of the time, against 76% to 79% for the other three quarters (fig. 7b), and the pattern is flatter for the report (67% in the bottom quarter and 74% to 78% elsewhere). Among the entries that flew, neither score is related to the flight score or to how close the flight came to its target altitude. A weak design and build score therefore goes with a higher chance of no flight, but a strong one doesn't say much about how good the flight will be.

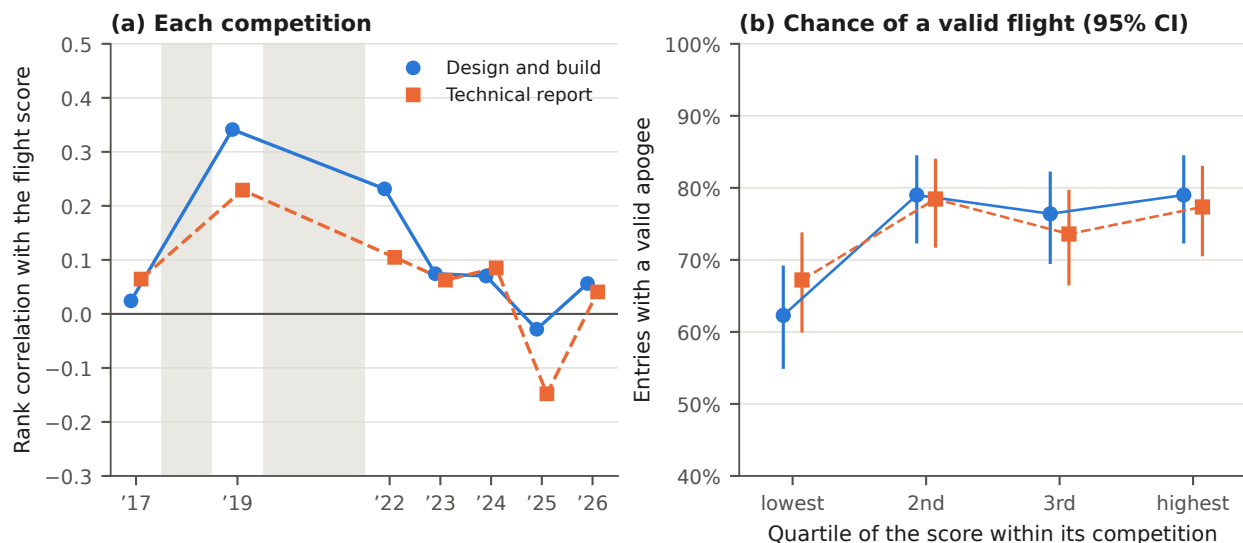


Figure 7: Rank correlation of each paperwork score with the flight score at each competition (a), and valid apogee rate by quartile of each score within its competition (b).

Table 1: Rank correlation of design and build scores and technical report scores with flight outcomes (95% intervals from resampling schools). The second row is the area under the ROC curve (0.5 means no information) and the last row gives standardised regression coefficients.

Comparison	Entries	Design and build	Technical report
Flight score, all entries	723	0.11 (0.03 to 0.18)	0.05 (−0.03 to 0.12)
Valid apogee recorded	723	0.60 (0.55 to 0.65)	0.55 (0.50 to 0.60)
Flight score, entries that flew	536	0.02 (−0.07 to 0.11)	0.01 (−0.08 to 0.09)
Altitude accuracy, entries that flew	536	0.02 (−0.07 to 0.10)	0.01 (−0.07 to 0.09)
Flight score, 2019 and 2022	200	0.29 (0.16 to 0.43)	0.17 (0.02 to 0.29)
Flight score, 2023–26	523	0.04 (−0.05 to 0.13)	0.00 (−0.09 to 0.09)
Both scores in one model	723	0.11 (0.03 to 0.19)	0.00 (−0.08 to 0.08)

The relationship was concentrated in the earlier years. The correlation between design and build and the flight score was 0.29 in 2019 and 2022 and 0.04 in 2023–26 (fig. 7a), and in the later years the design and build scores were also bunched closer together (the middle half of scores was 23 points wide in 2019 and 2022 and 15 points wide in 2023–26). A narrower range of scores lowers any correlation, so some fading is expected, but the data can’t say how much of it is that and how much is a change in how judges score. The shaded bands in the figure are 2018, which has no component breakdown, and the cancelled 2020 and virtual 2021 competitions. Two other points affect how this should be read. Design and build is judged at the competition from the poster session, discussions with the team, inspection of the hardware, and field preparation, so the judges see the rocket shortly before it flies, and a team whose hardware isn’t ready can both score low and not launch. The result is not evidence that a better design causes a better flight. The two paperwork scores are also related to each other (rank correlation 0.47), so the report isn’t independent of design and build, and its coefficient is about zero once design and build is accounted for.

Competitive balance at the top

Figure 8 shows (a) the gap between first and tenth place as a share of the winner’s total score and (b) the share of the field within 5% of the winner.

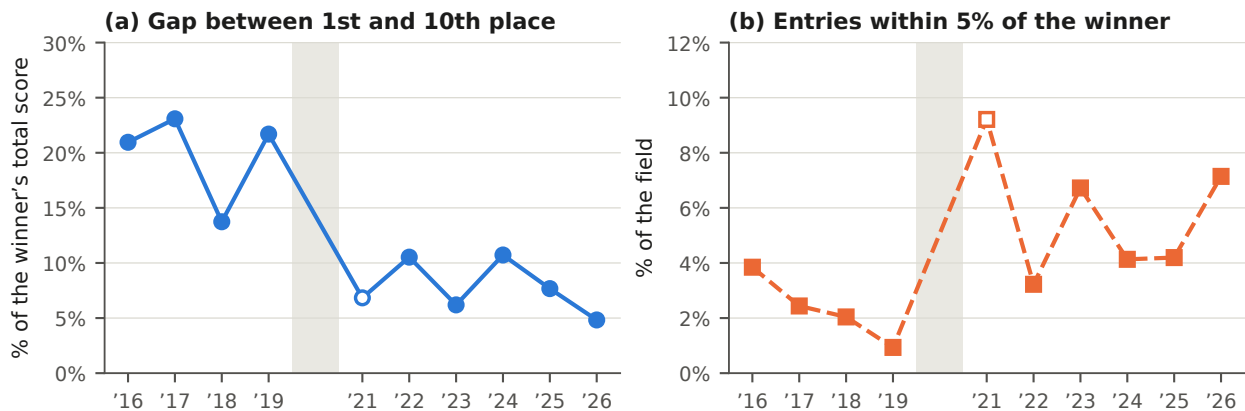


Figure 8: Gap between first and tenth place (a) and entries within 5% of the winner (b).

The gap between first and tenth was 14% to 23% of the winner’s score from 2016 to 2019 and was 4.8% in 2026. The share of the field within 5% of the winner went from 0.9% in 2019 to 7.1% in 2026, although it is noisy from year to year (2021 was 9.2%). The two highest total scores were 0.1% apart in 2023 (1.2 points) and 0.3% apart in 2025 (3.3 points). Part of the compression may be mechanical, since paperwork scores are capped and the field is larger, but combined with the variance results above it suggests that small differences in flight performance now decide the placings.

Comparing the classes

Figure 9 shows the median base score in (a) and the share of entries with a valid recorded apogee in (b) for each class, by competition. Classes at 10k are drawn with solid lines and circles and classes at 30k with dashed lines and squares, and cells with fewer than 5 entries are not plotted. The smaller classes have very few entries in any one year, so individual points are noisy, and the pooled before and after comparison later in this section is the better guide to what changed.

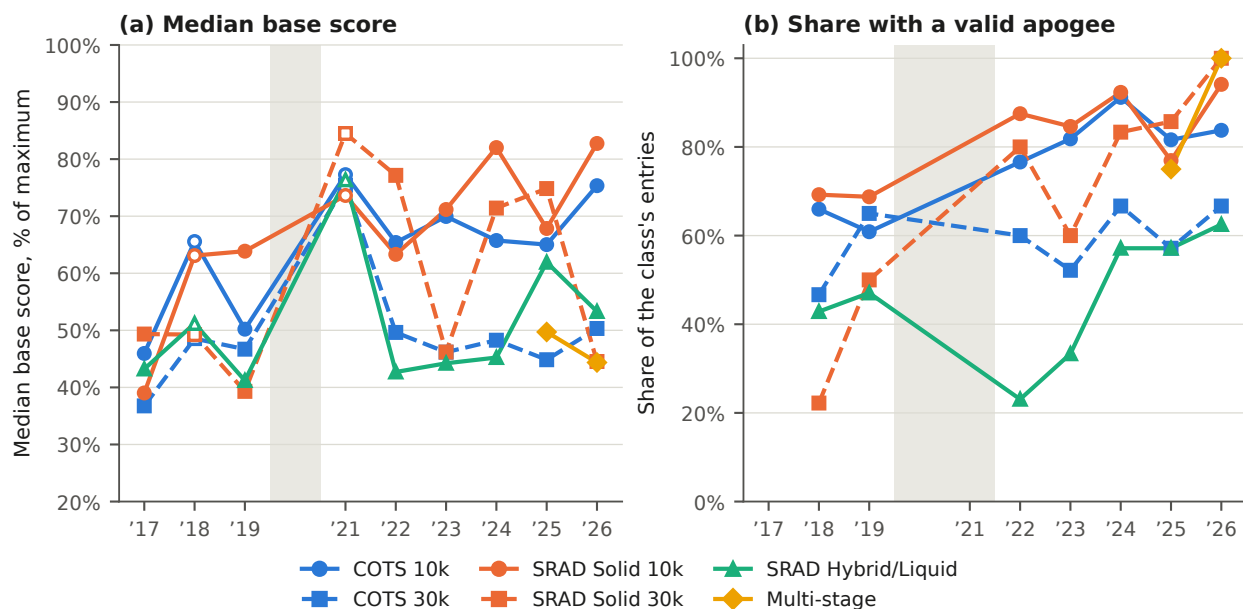


Figure 9: Median base score (a) and share of entries with a valid apogee (b), by class.

Reading across the two panels and table 2, 10k COTS is the broad base of the competition (56% of 2022–26 entries), with 83% of entries recording a valid apogee and a median altitude miss of 8%. 30k COTS is consistently harder, with 61% recording a valid apogee and a median miss of 18%. SRAD solid at 10k is now among the stronger classes (88% valid apogee, median standing 0.60). SRAD hybrid/liquid has the lowest completion rate, with 43% of 2022–26 entries recording a valid apogee (95% CI 30% to 57%), and those that flew missed the target by a median of 25%. Its median paperwork and design score over the same years was 91%, compared to 83% for 10k COTS. Multi-stage (14 entries in 2025 and 2026) is too new to say much about.

Table 2: Class profile, 2022–26 pooled. Recovered is among entries that flew; the last column is top-10 places, with the share of top-10 places against the share of entries.

Class	Entries	Valid apogee (95% CI)	Recovered	Median miss	Median standing	Top-10 places
COTS 10k	348	83% (79–87%)	74%	8.1%	0.56	27 (54% vs 56%)
COTS 30k	112	61% (51–69%)	62%	18.2%	0.33	11 (22% vs 18%)
SRAD Solid 10k	64	88% (77–94%)	79%	7.4%	0.60	6 (12% vs 10%)
SRAD Solid 30k	31	84% (67–93%)	50%	13.3%	0.48	4 (8% vs 5%)
SRAD Hybrid/Liquid	47	43% (30–57%)	70%	25.4%	0.29	2 (4% vs 8%)
Multi-stage	14	86% (60–96%)	33%	37.0%	0.30	0 (0% vs 2%)

To check which classes actually changed, fig. 10 pools 2017 and 2019 (before the pandemic gap) and 2022–26 (after it) and compares the medians. Confidence intervals come from resampling schools rather than entries, since entries from the same school aren’t independent. In (a), 30k COTS (+2.5 points, 95% CI -4.0 to $+11.5$) is the only class whose change can’t be told apart from zero. SRAD hybrid/liquid moved by a small but non-zero $+3.8$ (CI $+0.7$ to $+13.0$), and the gain for 30k SRAD solid is borderline (CI 0.0 to $+39.4$). Panel (b) shows the gap between pairs of classes. The 10k to 30k COTS gap widened from $+3$ to $+21$ points (change $+18$, CI $+6$ to $+28$). 10k COTS and 10k SRAD solid are statistically tied in both periods (recent gap -4 , CI -14 to $+4$). 10k SRAD solid now leads hybrid/liquid by $+28$ points (CI $+18$ to $+37$) compared to a tie earlier ($+1$), although the widening itself is borderline (CI -2 to $+38$).

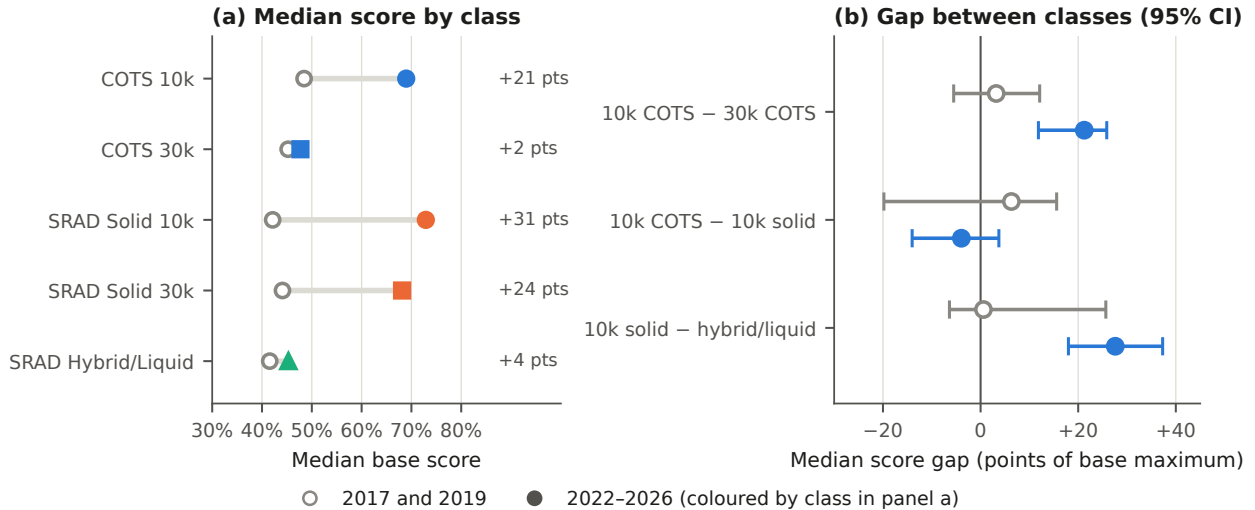


Figure 10: Median base score by class before and after the pandemic gap (a), and the gap between pairs of classes (b).

Table 3 gives the year by year trend for each class as a linear slope in percentage points per year, with school cluster bootstrap intervals. Mean score rose by a statistically significant amount in every class except 30k COTS ($+1.1$ per year, CI -0.3 to $+2.5$), and the all-class slope is $+2.2$ per year. The valid apogee share rose fastest for 30k SRAD solid ($+8.4$ per year, a small class) and not by a statistically significant amount for 30k COTS or hybrid/liquid. Measured as relative standing (the percentile in the field), none of the classes shows a significant linear drift, which is expected when most of the field is improving together, so the before and after gaps above are the sharper test. Hybrid/liquid has a positive mean slope alongside a small median change, which

would be consistent with its best teams improving more than its typical team, but I haven’t tested that directly.

Table 3: Linear trend by class in percentage points per year, with 95% cluster bootstrap intervals in brackets.

Class	Base score	Paperwork & design	Valid apogee	Standing
All classes	+2.2 [+1.7, +2.8]	+1.8 [+1.3, +2.3]	+3.3 [+2.3, +4.4]	0.0 [−0.8, +0.7]
COTS 10k	+2.3 [+1.5, +2.9]	+1.8 [+1.2, +2.5]	+2.9 [+1.4, +4.3]	−0.1 [−1.1, +0.8]
COTS 30k	+1.1 [−0.3, +2.5]	+1.3 [+0.2, +2.5]	+1.1 [−2.3, +4.0]	−1.0 [−3.0, +0.9]
SRAD Solid 10k	+3.1 [+1.7, +4.5]	+2.2 [+0.9, +3.9]	+2.9 [−0.2, +5.7]	+1.3 [−0.6, +3.1]
SRAD Solid 30k	+2.3 [+0.4, +4.1]	+2.4 [+0.3, +5.0]	+8.4 [+4.3, +12.3]	+0.4 [−2.0, +3.1]
SRAD Hybrid/Liquid	+2.2 [+1.0, +3.5]	+2.3 [+1.0, +4.0]	+1.4 [−3.1, +5.4]	−0.2 [−1.8, +1.6]

The single year gaps in fig. 10(b) are noisy, since the smaller class often has only 5 to 20 entries and most of the yearly intervals include zero. The 10k to 30k COTS gap was 4 to 17 points in 2017–19 and 16 to 25 points in every year since 2022. The 10k COTS and 10k SRAD solid gap sits on both sides of zero (SRAD solid was ahead by more than 5 points in 2019, 2024, 2026), and 10k SRAD solid was ahead of hybrid/liquid in 5 of the 5 competitions since 2022.

Finally, table 2 also shows whether any class wins more than its share of top ten places in 2022–26. The top ten places track class sizes closely (permutation test, $p = 0.58$), so no class is detectably over or under represented even though flight reliability differs a lot between them. With 50 places this test can only detect large differences.

School trajectories

Team numbers change from year to year, so the school is the unit used to follow trajectories. Figure 11 shows every school with seven or more appearances (41 schools). Each cell is the school’s percentile standing in that year and the letters give the class it entered (C for COTS, S for SRAD solid, H for hybrid/liquid, M for multi-stage, followed by the target altitude in thousands of feet; B and A are the 2016 Basic and Advanced classes). Rows are sorted by average standing, empty cells are years the school did not compete, and schools with more than one team in a year are represented by their best entry.

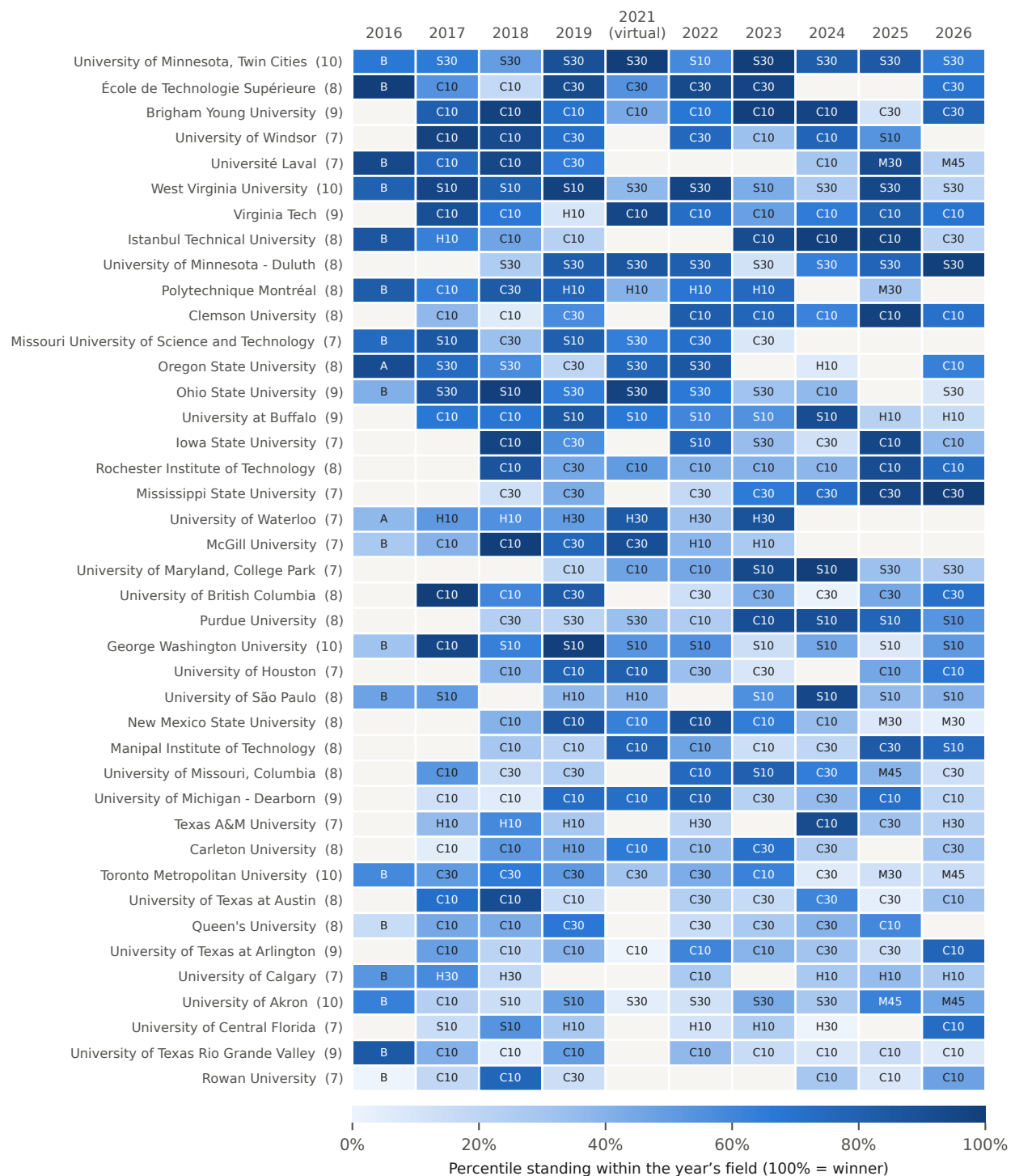


Figure 11: Percentile standing and class entered for schools with seven or more appearances (number of appearances in brackets).

A few schools are consistently dark (Minnesota (Twin Cities), ÉTS Montréal and BYU have the highest average standing), but most rows alternate between light and dark cells from one year to the next. Changes of class often coincide with a change in shade, which is covered in the next subsection.

To check whether individual schools have real trends, fig. 12 shows the schools with the steepest fitted slopes of standing against year, among the 79 schools with five or more appearances. 6 schools trend up and 5 trend down with 95% confidence, against roughly 2 in each direction expected by chance. So most single school trend lines can't be told apart from noise, and a school at the top or bottom of this list is partly selected for luck. The spread of slopes across schools is nonetheless larger than chance would give: the observed variance is 39.5, against 22.1 when each school's results are shuffled across years (95th percentile 28.5, permutation $p = 0.0005$). That implies a true spread of roughly ± 4.2 percentile points per year in schools' underlying trends. The average slope across schools is -0.22 per year, which is consistent with the field as a whole improving.

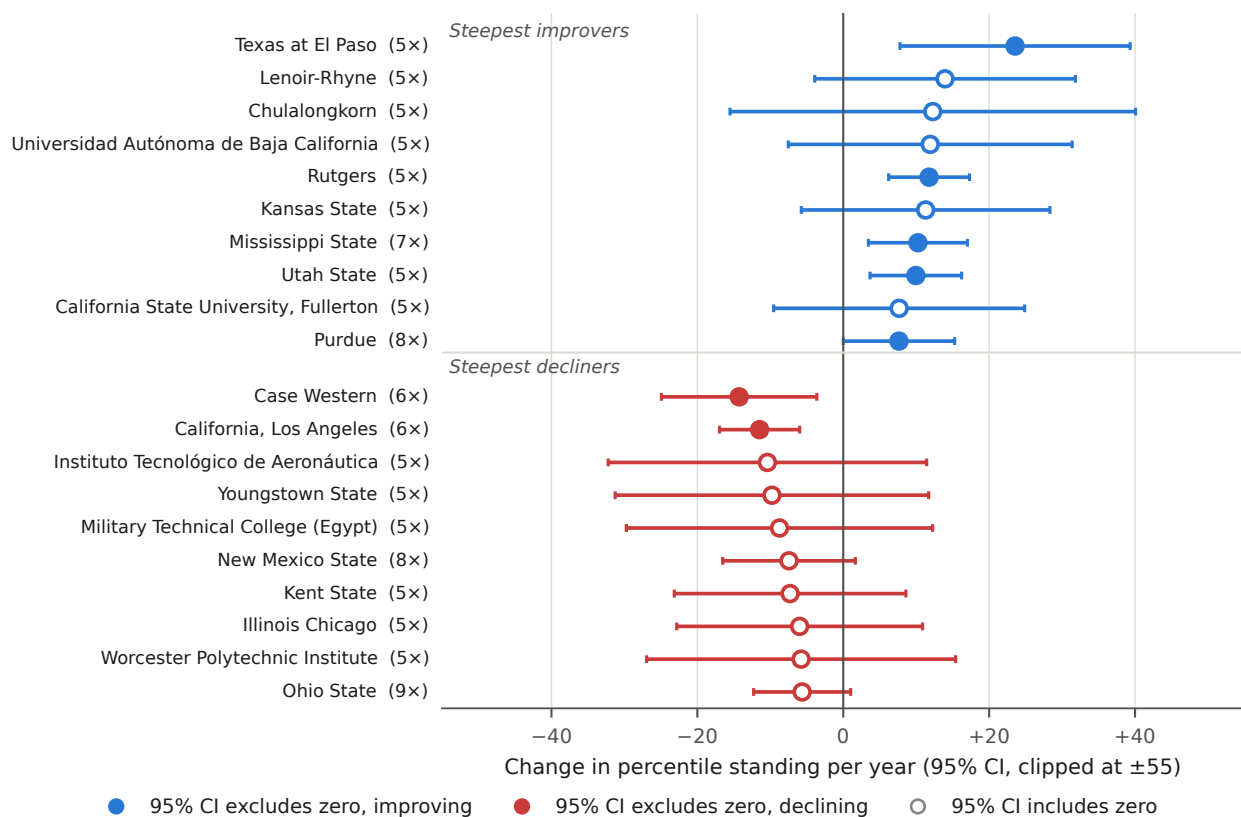


Figure 12: Schools with the steepest improving and declining trends in standing (five or more appearances).

Figure 13 has three related views. Panel (a) is the Spearman rank correlation between a school's result at one appearance and its next (excluding the pandemic transitions): 0.28 for overall standing, 0.46 for paperwork and design, and 0.28 for the flight score. Paperwork and design is a more stable trait of a school than flight results, which depend on single flights, weather, and range conditions. Panel (b) shows where schools finish at their next appearance depending on where they finished last time, split into fifths. The top fifth average the 65th percentile the next time and the bottom fifth average the 43th, so anyone comparing a good year with the next one should expect some

drop. Panel (c) is the mean standing by appearance number for schools first seen in 2018 or later (earlier schools may not have been new). The within school change per additional appearance is +1.6 percentile points (95% CI -0.6 to $+3.8$, 97 schools and 417 school-years), which suggests a small gain but is not distinguishable from zero.

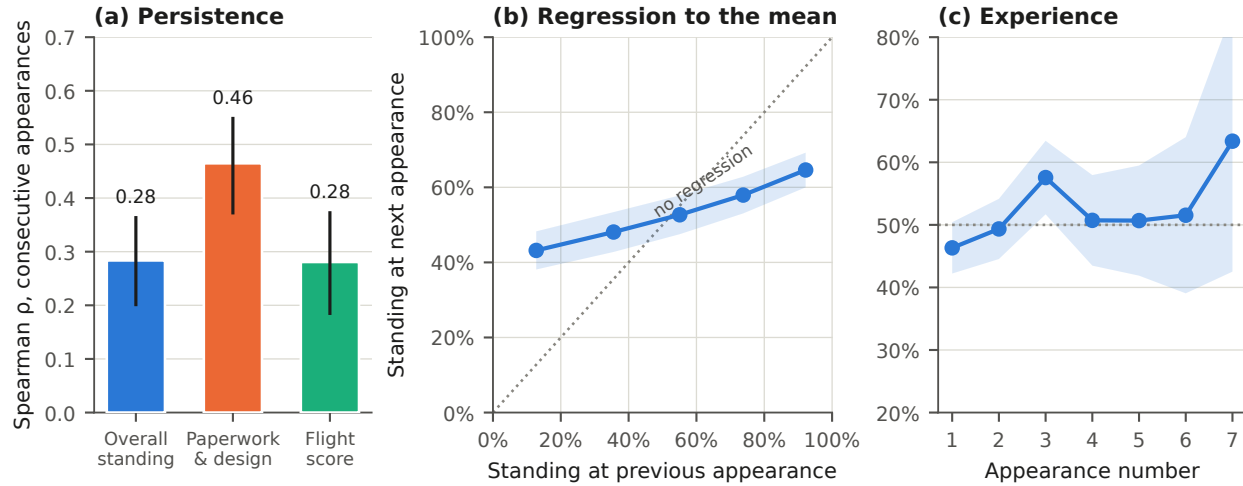


Figure 13: Year to year persistence (a), regression to the mean (b), and standing by appearance number (c).

Category switching

Schools often change class between appearances. A switch here means that a school's best entry at consecutive appearances is in a different class, where a step up is a higher target altitude or a move from COTS to SRAD, a step down is the reverse, and mixed means one of each. Figure 14 shows the counts for each pair of consecutive competitions, with the share of returning schools that changed class printed above each bar (547 school pairs in total). The pairs 2019 to 2021 and 2021 to 2022 span the virtual competition. Overall 28% of pairs involve a change of class, 19% change altitude, and 14% change propulsion type. Switching dipped to 15% across the 2021 to 2022 return, which is expected, since the virtual 2021 competition gave teams no flight result, good or bad, to react to.

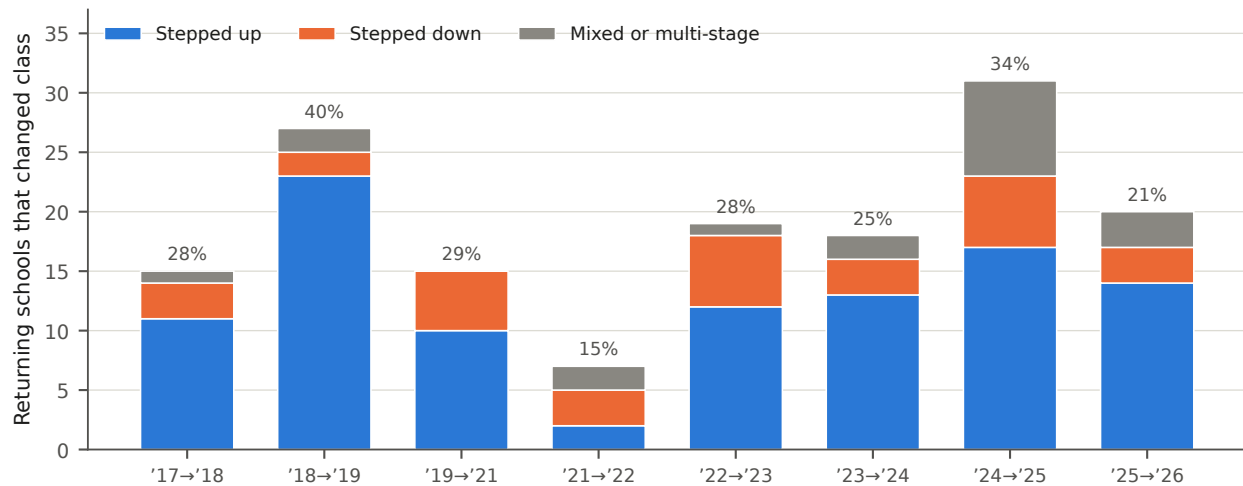


Figure 14: Class changes between consecutive appearances.

In fig. 15(a), each row is the class a school entered and each column the class it entered at its next appearance, with the counts printed in the cells. Staying in the same class is the norm (72% overall), most strongly for hybrid/liquid (87% stay) and least for 10k SRAD solid (35% change). 10k COTS teams that move mostly go to 30k COTS (46 moves) rather than to SRAD (35 moves). Panel (b) shows what happens to a school's standing after each kind of move. The raw change for schools stepping up is -17.3 percentile points, but these schools were already strong (average standing 0.64 before the move), so some of the drop is regression to the mean. After adjusting for that, using the relationship fitted on schools that stayed in their class, a step up costs -12.0 points (95% CI -17.5 to -6.5). The apparent gain from stepping down ($+13.4$ raw) shrinks to $+5.0$ after the same adjustment (CI -4.0 to $+14.4$).

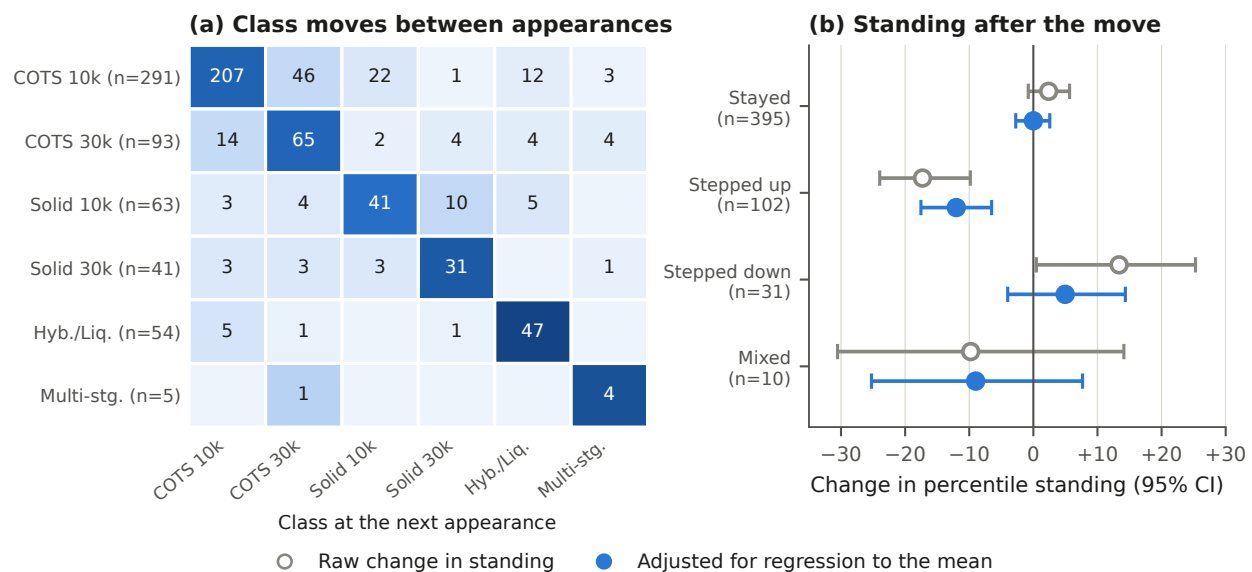


Figure 15: Class moves between appearances (a) and change in standing after each kind of move (b).

Figure 16 compares the probability of each move at the next appearance for schools that did and did not record a valid apogee in the current year (flight years only). Schools without a valid apogee were more likely to retreat: from SRAD to COTS (20% against 6%, $p = 0.028$) and from 30k or higher to 10k (26% against 12%, $p = 0.051$). Schools that flew were more likely to climb from 10k COTS to 30k or higher (21% against 7%, $p = 0.047$). Moving from COTS to SRAD showed no clear difference ($p = 0.407$). Three of the four tests point the same way, but each is borderline. These are four tests on the same data, so I also adjusted them for each other (Holm's method), which gives $p = 0.11, 0.14, 0.14$, and 0.41 . None is below 0.05 after the adjustment, so I'd treat the pattern as suggestive.

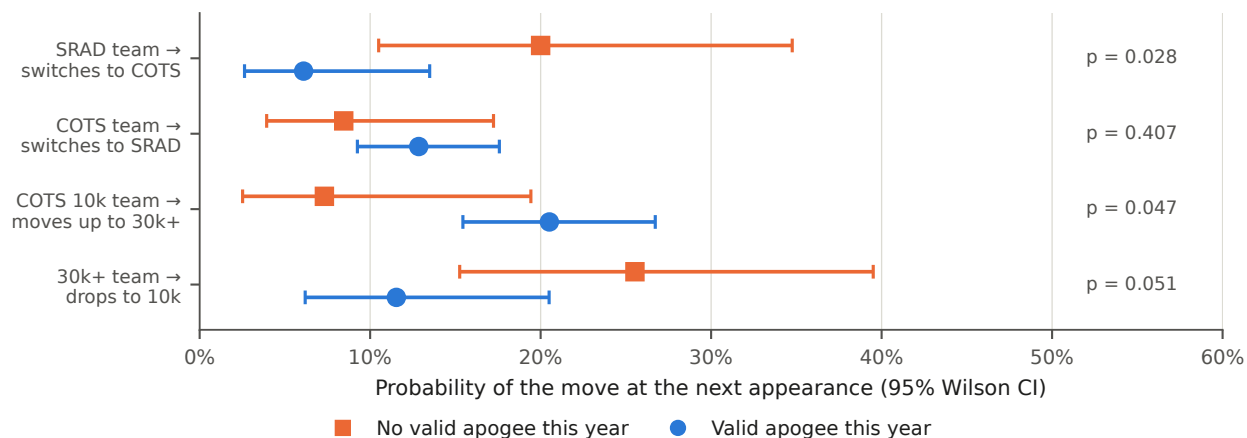


Figure 16: Probability of changing class at the next appearance, by whether the school recorded a valid apogee.

Figure 17 shows how many classes schools try. Of the 112 schools with four or more appearances, 54 used two different classes and 24 stayed in one, and the average was 2.2. Of the 73 of those schools that started in 10k COTS, about two thirds eventually tried a 30k or higher class, about half tried SRAD motors, and 29% tried both.

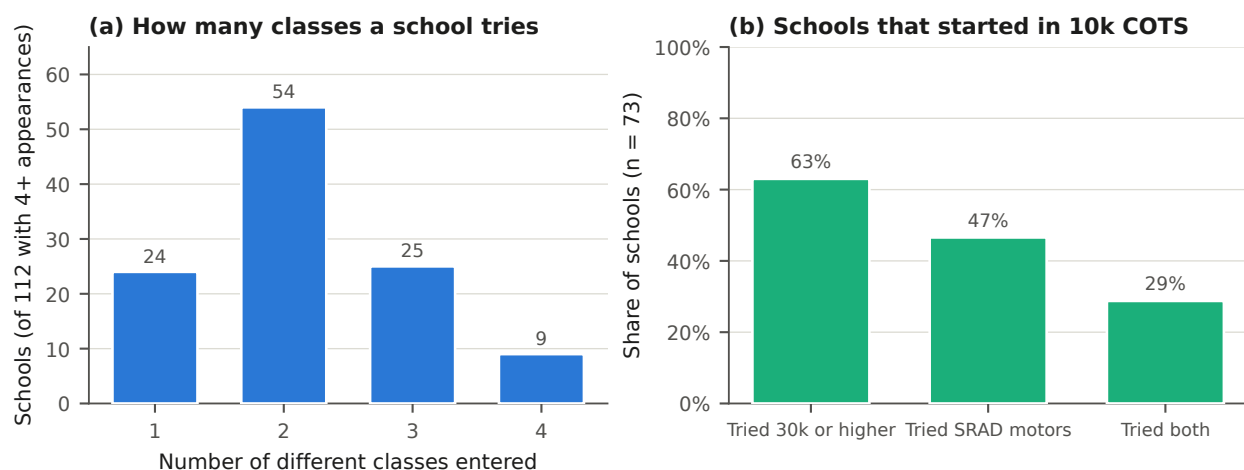


Figure 17: Number of classes entered by schools with four or more appearances (a), and what schools that started in 10k COTS went on to try (b).

Flight performance

Figure 18 looks at how close flights come to their target altitude, among teams that recorded a valid apogee (2018, 2019, and 2022–26). Panel (a) shows the median miss (the absolute difference between apogee and target, as a share of the target). Panel (b) shows the share of flights within 10% of the target and the share that missed by 30% or more, which earns zero altitude points under the 2022 and later formula. Panel (c) shows the median signed error, where a negative value is a flight that finished below its target, and panel (d) shows the share of flights that finished above the target.

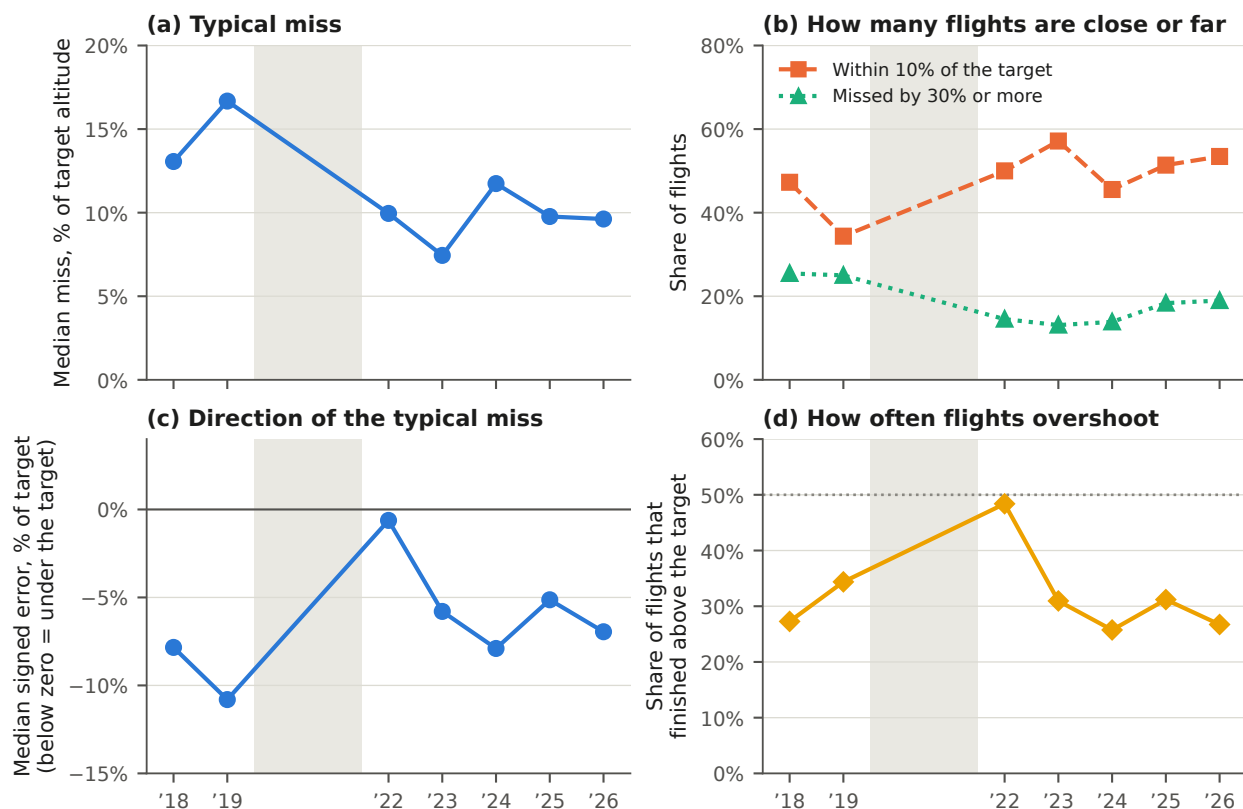


Figure 18: Altitude accuracy (a) and direction of the miss (b), among flights with a valid apogee.

The median miss was 13% in 2018 and 17% in 2019, and has been between 7% and 12% in 2022–26. The trend is -0.7 points per year (95% CI -1.5 to 0.0), which is borderline. The share missing by 30% or more fell from about a quarter in 2018 and 2019 (25% and 25%) to 13% in 2023 and has since risen to 19% in 2026. I can't tell from this data why. Altitude is also harder to hit at height: among 2022–26 flights the median miss is 9% at 10k and 18% at 30k and above.

Flights consistently finish below target. The median signed error is negative in every year, between 0.6% and 10.8% below target, and only 26% to 48% of flights overshoot in any year, with 2022 closest to balanced (panels c and d).

Figure 19 shows how well teams predicted their own apogee. The median error went from 12% in 2018 to 9% in 2025, but not steadily (slope -0.5 points per year, CI -1.4 to $+0.4$, which includes zero), and the share of flights within 5% of the prediction went from 25% in 2018 to 34% in 2025 (it was 14% in 2019). The median flight in 2025 finished 5.6% below its prediction. The 2026 results don't include predictions.

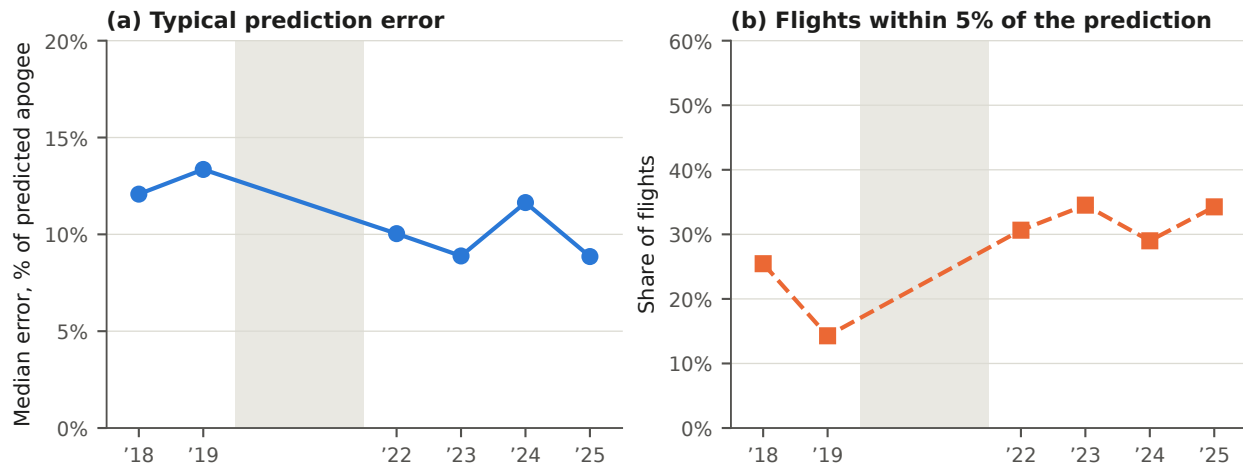


Figure 19: Median prediction error (a) and share of flights within 5% of the team's predicted apogee (b).

The CubeSat bonus

Teams whose functional payload adopts the CubeSat form factor earn a 50 point bonus, which is 5% of the base maximum. A payload that is only ballast has to use the CubeSat form anyway, so the bonus rewards building a functional payload in that form. The score sheets list it separately in 2019 and 2022–25, and the 2026 PDF has a single bonus column whose only values are 0 and 50, which I read as this bonus. The 2018 file reports totals only and the 2017 scoring had no payload bonus, so the comparison starts in 2019. Figure 20 shows (a) the share of entries earning the bonus in each competition and (b) the median standing of entries with and without it. Standing here is the percentile on the base score (every bonus removed), so the 50 points do not move anyone up the table. The shaded band covers the cancelled 2020 and virtual 2021 competitions.

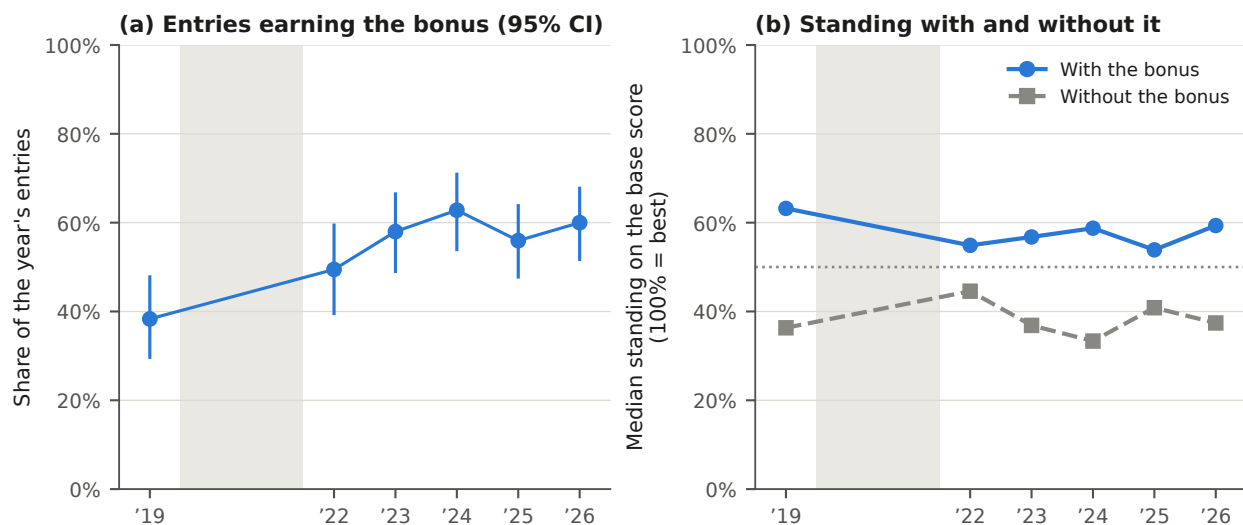


Figure 20: Entries earning the CubeSat bonus (a), and median standing on the base score with and without it (b).

The share of entries earning the bonus rose from 38% in 2019 to 49% in 2022 and has been between 56% and 63% in every competition since 2023. It is more common among SRAD entries (about three quarters of SRAD solid and hybrid/liquid entries in 2022–26) than among COTS entries (about half), and about as common at 10k as at 30k and above.

Table 4: Entries with and without the CubeSat bonus, 2019 and 2022–2026 pooled. Standing and top ten places use the base score, which excludes all bonuses.

	With the bonus	Without
Entries	396	327
Median base score	69%	51%
Median standing on the base score	58%	38%
Valid apogee recorded	81%	65%
Recovered, of the entries that flew	70%	65%
Median altitude miss, of the entries that flew	9.5%	12.1%
Median paperwork and design score	86%	81%
Top ten places on the base score (of 60)	48	12

Entries with the bonus stand higher on the base score in every competition (fig. 20b). Pooled over the competitions, the mean difference in standing is 14 percentile points (95% CI 10 to 19), and 15 points when entries are compared only within their own class and year. They also recorded a valid apogee more often, and 48 of the 60 top ten places on the base score went to entries with the bonus, which were 55% of all entries (table 4). Among the 105 schools that appear in some years with the bonus and some without, a school’s standing was 9 points higher in the years with it (95% CI 3 to 15), which is smaller than the gap between teams. Part of the gap is therefore probably about which teams build a functional payload in the CubeSat form, and the data can’t say whether the payload itself helped.

Overall winners and geography

The overall winner isn’t simply the highest score. The Rules and Requirements say that one of the six category first-place teams is named overall winner by the competition officials, using qualitative assessments (2022 draft rules) or quantitative and qualitative assessments (2027 rules), and the organizers’ past-events page describes the 2017 and 2019 choices as “chosen qualitatively” from the category winners. Table 5 therefore gives the declared winner, where that entry ranked on total score, and the highest-scoring entry when it was a different team.

There have been nine declared winners from 2017 to 2026, from eight schools. The University of Sydney won in 2022 and 2025 and no other school has won more than once, so no school has won at consecutive competitions. The declared winner had the highest total score in five of the nine years. In 2023 it was the second highest (BYU, 1.2 points behind Minnesota) and in 2025 the third highest (Sydney, 10 points behind Virginia), and in 2017 and 2019 it ranked 33rd of 82 and 24th of 107. The page gives no basis for the 2023 and 2025 choices. US schools have won six of the nine titles, Australia two (both Sydney), and Canada one (McGill, 2018). The highest total score belonged to a Canadian team in each of 2016 to 2018, the first three competitions in this dataset, but 2016 was not the first IREC, and only McGill in 2018 was also the declared winner. The 2016 results had separate Basic and Advanced winners and no overall title, and the 2016 total of 1,281 used a different 1,400 point rubric and isn’t comparable to the others.

Table 5: Overall winner declared by the organizers at each competition, where it ranked on total score, and the highest-scoring entry when that was a different team.

Year	Declared winner	Country	Class	Total	Rank by total	Highest total score
2016	No overall winner declared. Basic: École de Technologie Supérieure; Advanced: Oregon State University.					
2017	University of Michigan, Ann Arbor	USA	30k hybrid	459	33 of 82	UBC (863)
2018	McGill University	Canada	10k COTS	1,103	1 of 98	same
2019	University of Washington	USA	30k hybrid	795	24 of 107	Sydney (1,147)
2021	University of Minnesota, Twin Cities	USA	30k solid	956	1 of 76	same
2022	University of Sydney	Australia	30k COTS	1,046	1 of 93	same
2023	Brigham Young University	USA	10k COTS	1,117	2 of 119	Minnesota (Twin Cities) (1,118)
2024	University of Maryland, College Park	USA	10k solid	1,106	1 of 121	same
2025	University of Sydney	Australia	10k hybrid	1,034	3 of 143	Virginia (1,044)
2026	Mississippi State University	USA	30k COTS	1,002	1 of 140	same

Table 6 lists the schools that have reached the top ten at least three times. 67 different schools have reached a top ten across the ten competitions and 23 of them did so more than once, but nobody has done it more than 4 times. The University of Sydney reached the top ten in 4 of its 5 appearances.

Table 6: Schools with three or more top-ten finishes (years in bold are declared overall wins; none was declared in 2016).

School	Country	Top-10 finishes	Appearances	Top-10 years
University of Sydney	Australia	4	5	2019, 2021, 2022 , 2025
École de Technologie Supérieure	Canada	4	8	2016, 2019, 2022, 2023
Istanbul Technical University	Turkey	4	8	2016, 2023, 2024, 2025
West Virginia University	USA	4	10	2017, 2019, 2022, 2025
Brigham Young University	USA	3	9	2018, 2023 , 2024

For geography, table 7 compares the periods before (2016–2019) and after (2022–2026) the pandemic gap, with the virtual 2021 competition excluded. Each cell shows the earlier value then the later one, and the last column is the p -value of a Mann–Whitney test on standing between the two periods. Canada is the clearest change: entries per competition went from 16 to 10, median standing from 0.57 to 0.37, and its share of top ten places from 32% to 6% ($p = 0.004$, or 0.03 after a Bonferroni adjustment for the six comparisons). It is the only region whose change in standing is statistically clear. The fall in Canadian entries most likely reflects Launch Canada (see the field subsection), and the data can’t say whether the fall in standing is related to it. Oceania stands out in 2022–26: 25 entries (4% of the field, 22 of them Australian) with a median standing of 0.75 and 7 of the 50 top ten places, and both of the declared winners from the region were Sydney. Turkey’s growth is in volume: its median standing in 2022–26 (0.46) is close to the field’s and not better.

Table 7: Region comparison, 2016–2019 → 2022–2026 (2021 excluded). Wins are declared overall wins, and none was declared in 2016. Oceania had one entry before the pandemic, so its median standing and p -value are not given.

Region	Entries / yr	Share of field	Median standing	Top-10 share	Declared wins	p
USA	53.8 → 75.6	63% → 61%	0.49 → 0.52	52% → 64%	2 → 3	0.489
Canada	16.0 → 10.0	19% → 8%	0.57 → 0.37	32% → 6%	1 → 0	0.004
Latin America	4.0 → 8.0	5% → 6%	0.41 → 0.44	2% → 4%	0 → 0	0.543
Europe	4.2 → 6.6	5% → 5%	0.48 → 0.32	5% → 6%	0 → 0	0.436
Turkey	2.5 → 8.8	3% → 7%	0.29 → 0.46	2% → 6%	0 → 0	0.442
Oceania	0.2 → 5.0	0% → 4%	– → 0.75	2% → 14%	0 → 2	–
Asia, Middle East and Africa	4.0 → 9.2	5% → 7%	0.27 → 0.43	2% → 0%	0 → 0	0.647

Category winners

A category winner here is the organizers’ first place in an official category (target altitude by propulsion type, or Basic and Advanced in 2016), as listed on the ESRA past-events page and matched to the score files by team number. Where the page lists no winner (the 2016 results, which have no team numbers and agree with the highest scores, and three 2026 categories with no winner), the highest total score in the category is used. A category-year with fewer than 5 entries is called thin. There are 12 thin category-years out of 60, and they’re shown in fig. 21 but excluded from the statistics below.

The listed winner is also the highest-scoring entry in 50 of the 55 category-years that have a listed first place (the six 2021 virtual-competition entries are labelled honorable mentions on the page). In four of the other five, the higher-scoring entry doesn’t appear to have been eligible for the place award. The 2017 score sheet marks Istanbul Technical University “Not Eligible for Place” in 10k hybrid/liquid, and the higher-scoring entries in 2018 (Oregon State in 30k solid, Michigan in 10k hybrid/liquid) and 2024 (Wrocław Tech in 10k hybrid/liquid) either have no valid apogee or reached 3,864 ft against a 10,000 ft target, whereas the 2022 draft rules require a successful launch to at least half the target altitude. The fifth, 10k hybrid/liquid in 2023, is listed with the same winner as 10k solid (Cornell), which doesn’t match the score files, so I kept the highest score there (Polytechnique Montréal). The 2023 30k COTS category is listed as shared between ÉTS Montréal and Canterbury, and Canterbury, which scored higher, is shown. Cells that differ from the highest score are marked in fig. 21.

Figure 21 lists the winner of every category at every competition. The thin coloured bar on the top of a cell shows the winner’s region, a star marks the declared overall winner, and a circular arrow marks a school that also won the same category at the previous competition. 60 category titles have gone to 42 schools from 8 countries, 48 of them in categories with 5 or more entries. University of Sydney is the most versatile winner, with titles in 3 categories (10k COTS, 10k hybrid, 30k COTS), and University of Minnesota, Twin Cities is the most dominant in a single class (4 titles in 30k solid). Only 3 schools have won in more than one category, and 70% of titles (42 of 60) went to a school that had never won one before. Winners are not simply the most experienced schools: on average they had +0.26 more prior appearances than the rest of their category-year (95% CI −0.21 to +0.74 across 35 category-years).

	2016	2017	2018	2019	2021 (virtual)	2022	2023	2024	2025	2026
10k COTS		UBC n=45	McGill n=47 ★	Sydney n=46	Virginia Tech n=36	Case Western n=47	BYU n=66 ★	Istanbul Technical n=68	Virginia n=87	UC Santa Cruz n=80
30k COTS		Stanford n=6	Concordia n=15	Georgia Tech n=20	NTNU n=8	Sydney n=20 ★	Canterbury n=23	Canterbury n=27	Mississippi State n=21	Mississippi State n=21 ★
10k SRAD Solid		West Virginia n=12	Ohio State n=13	George Washington n=16	Buffalo n=7	Kent State n=8	Cornell n=13	Maryland n=13 ★	Fed. U. of ABC n=13	Texas Tech n=17
30k SRAD Solid		Ohio State n=7	Illinois Chicago n=9 †	Minnesota (Twin Cities) n=8	Minnesota (Twin Cities) n=11 ★	West Virginia n=5	Minnesota (Twin Cities) n=5	Minnesota (Twin Cities) n=6	West Virginia n=7	Minnesota Duluth n=8
10k SRAD Hybrid/Liquid		Waterloo n=7 †	Waterloo n=11	Poly Montréal n=13	AGH Kraków n=10	Poly Montréal n=8	Poly Montréal n=8	Calgary n=4 †	Sydney n=6 ★	Sydney n=7
30k SRAD Hybrid/Liquid		Michigan n=5 ★	Washington n=3	Washington n=4 ★	Poznań Tech n=4	Tennessee n=5	Poznań Tech n=4	Rice n=3	Rice n=1	Texas A&M n=1
30k Multi-stage									Laval n=4	Baylor n=2
45k Multi-stage									Akron n=4	Akron n=4
Basic (2016)	ÉTS Montréal n=44									
Advanced (2016)	Oregon State n=8									

USA Latin America Turkey Asia, Mid. East, Africa repeat winner organizers' listing unusable; top scorer shown
 Canada Europe Oceania ★ declared overall winner † listed winner, not the top scorer n italic grey: fewer than 5 entries

Figure 21: Category winner at every competition (n is the number of entries in the category).

Figure 22 compares how often winners repeat with what would happen if each year’s winner were drawn at random from that category’s entries. The same school won a category at consecutive competitions in 10 of 50 cases, where chance predicts about 3 (simulation, $p < 0.001$). The repeats are marked in fig. 21, and Minnesota (Twin Cities) accounts for two of them in 30k solid (2019 to 2021 and 2023 to 2024).

Of the 41 contested category winners that had a later competition, 36 came back and 25 of those returned to the same category. Seven won again and 16 finished in the category top three, where chance would give about 2 and 6. Returning winners average the 70th percentile in the whole field next time. This sits alongside the weak persistence for the field as a whole ($\rho = 0.28$), which suggests that persistence is concentrated at the top. No school has ever won the overall title at consecutive competitions.

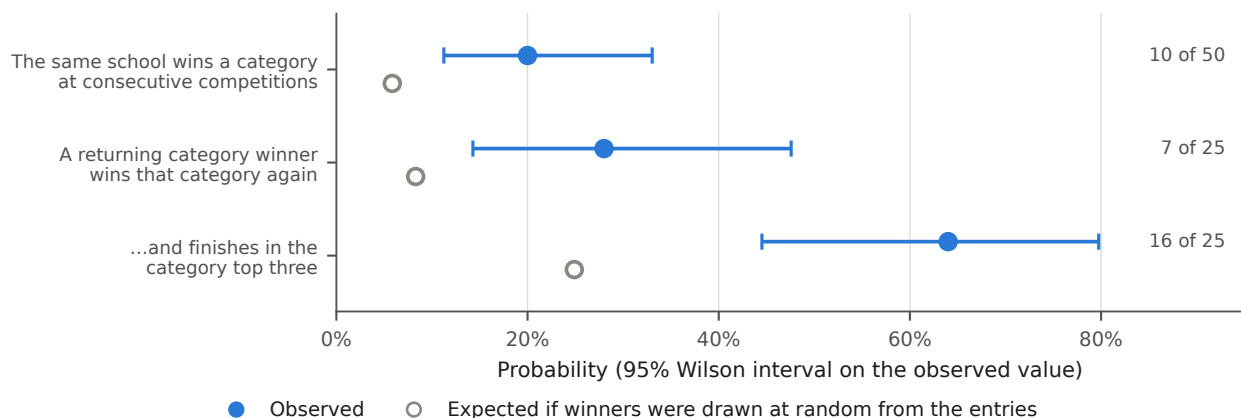


Figure 22: Category winners repeating, compared with random draws from the entries.

Figure 23 shows (a) the winning base score in each category and (b) the margin of victory as a share of the winner’s total, for categories with at least 5 entries. Winning scores rose from 65% in 2017 to 85% in 2019 and have been near 91% since 2022 (88% to 93% in any year). Excluding 2017, the pooled trend is +0.7 points per year (95% CI -0.4 to $+1.7$), which is flat and consistent with winners in every class approaching the same ceiling.

The median margin of victory is 4.0% across the 43 contested category-years where the winner was also the top scorer, with 10 decided by under 1% and 20 by more than 5%. By comparison, the two highest total scores in a year were a median of 1.1% apart. 30k COTS is the tightest category (median 1.9%) and 10k hybrid/liquid the loosest (23.9%), although the fields there are small.

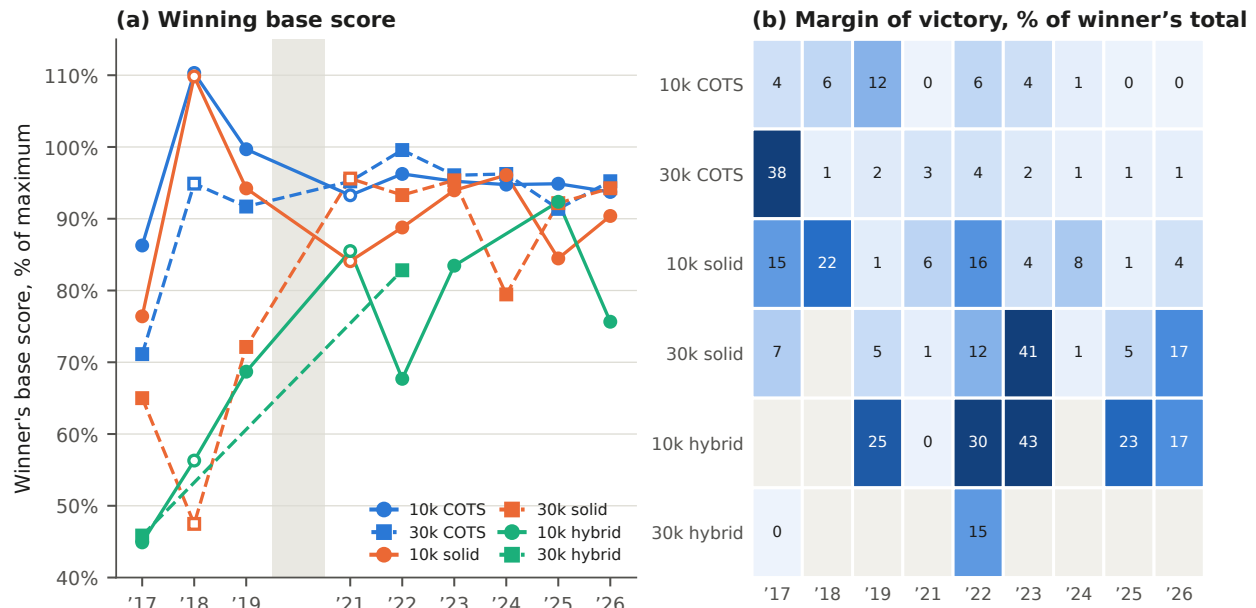


Figure 23: Winning base score (a) and margin of victory in percent (b), by category.

Figure 24 compares the 25 category winners in contested categories since 2022 with all other entries in those categories. Every one of the winners recorded a valid apogee and recovered, against about three quarters and a little over half of the rest of the field. The winners' median altitude miss was 1.9% compared to 9.5%, and 80% were within 5% of target compared to 32%. Paperwork and design scores are higher for winners (median within-category percentile 0.83), but flight is what separates them most (0.99). Winners also incurred penalties less often (28% of entries compared to 43%).

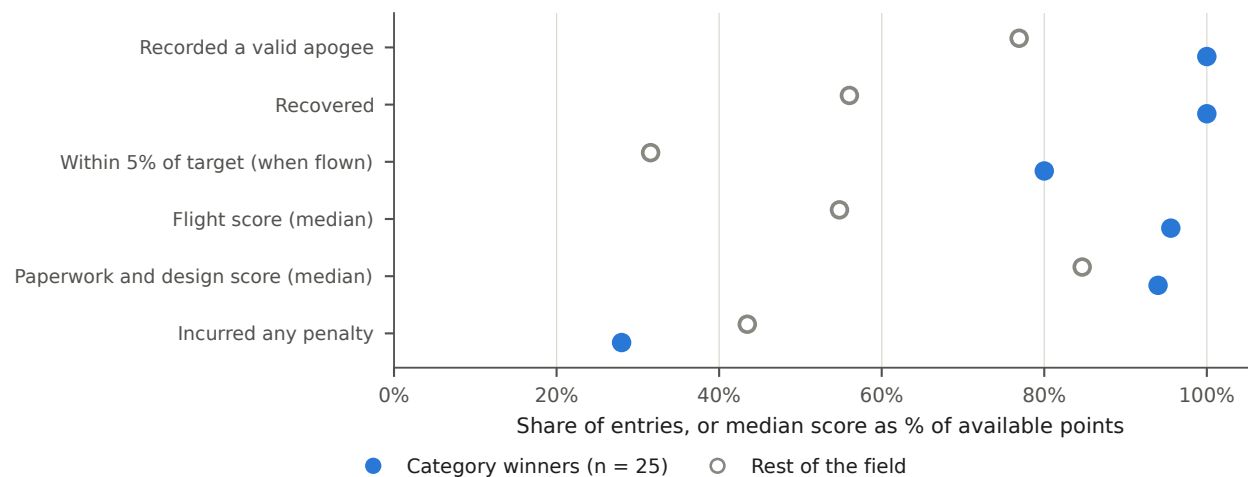


Figure 24: Category winners compared with the rest of the field in contested categories, 2022–2026.

Table 8 lists the schools with the most category titles. Of the 42 schools with a category title, 4 have won three or more and 34 have won exactly once; the table lists those with two or more. Polytechnique Montréal's 2023 title is the highest score in the category whose organizer listing

was unusable. Table 9 compares each region’s share of category wins with its share of entries, for contested categories only. In 2022–26 the shares line up for every region except Oceania, which won 5 of the 25 contested categories with 4% of the entries. In 2017–19 Canada won more than its share of entries (6 of 16 against 18%), and its wins then went from 6 to 2 as its entries fell. These are small counts, so I would treat them as descriptive.

Table 8: Schools with two or more category titles.

School	Country	Titles	Categories won	Appearances	Years won
Minnesota (Twin Cities)	USA	4	30k solid	10	2019, 2021, 2023, 2024
Sydney	Australia	4	10k COTS, 10k hybrid, 30k COTS	5	2019, 2022, 2025, 2026
Poly Montréal	Canada	3	10k hybrid	8	2019, 2022, 2023
West Virginia	USA	3	10k solid, 30k solid	10	2017, 2022, 2025
Mississippi State	USA	2	30k COTS	7	2025, 2026
Ohio State	USA	2	10k solid, 30k solid	9	2017, 2018
Canterbury	New Zealand	2	30k COTS	3	2023, 2024
Waterloo	Canada	2	10k hybrid	7	2017, 2018

Table 9: Category wins by region (contested categories). Shares are of wins versus of entries.

Region	Wins 2017–19	Share of wins vs entries	Wins 2022–26	Share of wins vs entries
USA	9 of 16	56% vs 64%	16 of 25	64% vs 62%
Canada	6 of 16	38% vs 18%	2 of 25	8% vs 7%
Latin America	0 of 16	0% vs 5%	1 of 25	4% vs 6%
Europe	0 of 16	0% vs 5%	0 of 25	0% vs 5%
Turkey	0 of 16	0% vs 3%	1 of 25	4% vs 7%
Oceania	1 of 16	6% vs 0%	5 of 25	20% vs 4%
Asia, Middle East and Africa	0 of 16	0% vs 5%	0 of 25	0% vs 8%

Conclusions

This section collects the results that the data support well, meaning the change is large compared to its uncertainty and shows up across most competitions, followed by the results that are weaker. Table 10 gives the average change per year in the main measures. Each slope rests on seven or fewer competitions, so it describes direction and rough size and is not a forecast.

Table 10: Average change per year in the main measures, in percentage points of the available points or of the entries, with 95% cluster bootstrap intervals. The score rows use 2017, 2019, and 2022–26, valid apogee and altitude miss use 2018, 2019, and 2022–26, and the recovery rows use 2019 and 2022–26.

Measure	Change per year	95% interval
Base score, all entries	+2.2	+1.7 to +2.8
Admin points	+1.5	+1.0 to +2.1
Technical report	+1.3	+0.7 to +2.0
Design and build	+2.4	+1.8 to +3.1
Flight score, all entries	+3.5	+2.6 to +4.3
Valid apogee recorded	+3.3	+2.3 to +4.4
Recovered, of the entries that flew	+2.6	+0.7 to +4.5
Valid apogee and recovery, all entries	+4.0	+2.4 to +5.6
Altitude miss, of the entries that flew	−0.7	−1.5 to 0.0

Who competes. Entries grew from 82 in 2017 to 140 in 2026, and most of the net increase is COTS, mainly 10k COTS (35 of the net 58, with 15 more in 30k COTS). Most new schools start there as well: 78% of the schools first seen from 2018 entered in 10k COTS. The US share of entries has stayed at about 60%, so US entries grew with the field. Turkish entries grew about 3.5 times (from 2.5 to 8.8 per competition) and Canadian entries fell from 16 to 10, with Canada’s share going from 22% in 2017 to 4% in 2026. The other regions each grew by a factor of 1.6 to 2.3, and Oceania went from a single entry before the pandemic to about 5 per competition.

Scores. The median base score rose about 2.2 points per year, from 44% in 2017 to 72% in 2026, and the rise is in the middle and bottom of the field. The top has not moved since 2022 (the 90th percentile at 88% to 90% and the winning score at 95% to 100%), so the gap between first and tenth place fell from 22% of the winner’s score in 2019 to under 5% in 2026.

Components and report scoring. Every component has risen on average, but not in the same way. The median entry has earned full admin points in every year since 2017, so the change there is at the bottom of the field. The technical report is similar: its median has been flat at 80% to 86% since 2022 while the weakest reports improved. Design and build and flight scores rose most, and design and build scores also became much less spread out (the middle half of scores was 31 points wide in 2017 and 12 in 2026).

Flight and recovery. The share of entries that recorded a valid apogee went from 56% in 2018 to 83% in 2026, and recovery among the entries that flew went from 52% to 71%, less steadily. Together, the share of entries with both a valid apogee and a recovery nearly doubled, from 31% in 2019 to 59% in 2026. Altitude accuracy improved only slightly (the median miss was 17% in 2019 and 7% to 12% since 2022), and flights have finished below the target in every year.

What decides the total score. Flight accounts for a larger share of the variance in total score (51% in 2019 and 65% to 79% in 2022–26), while the report and design and build together went from 35% in 2017 to 10% in 2026. Part of the shift is mechanical, since paperwork scores have bunched near their caps and so separate teams less. Bonus minus penalty has been the second largest source of variance in every competition since 2023 (18% to 25% in 2023–25 and 11% in 2026, when the only bonus in the file is the 50 point CubeSat bonus), ahead of the report and design and build.

Classes. The median base score rose by 21 to 31 points in 10k COTS, 10k SRAD solid, and 30k SRAD solid, but not detectably in 30k COTS, so 10k COTS now leads 30k COTS by 21 points

where it led by 3 before the pandemic. Returning schools mostly move to harder classes: 102 of the 152 class changes were a step up and 31 were a step down, and step-ups outnumber step-downs in every pair of consecutive competitions except 2021 to 2022, the first competition after the virtual one, when teams had no flight result to react to. After adjusting for regression to the mean, a step up costs about 12 percentile points of standing.

Schools. Paperwork and design is the most persistent school trait. The rank correlation between a school’s score at one appearance and its next is 0.46 (95% CI 0.37 to 0.55), against 0.28 (0.18 to 0.37) for the flight score and 0.28 (0.20 to 0.36) for overall standing. That persistence has not turned into a measurable gain with experience (+1.6 percentile points per appearance, 95% CI -0.6 to $+3.8$). Schools also come back less often than before the pandemic: 78% of one competition’s schools appeared at the next in 2016–19 and 69% in 2022–26 (Fisher exact $p = 0.019$), and schools in their first appearance are the least likely to return.

Winners. The declared overall winner is chosen from the category winners with qualitative judgement and had the highest total score in five of nine years, so the score alone doesn’t predict it. Category titles repeat more often than chance (10 of 50 cases against about 3 expected), and the category winners are the teams that fly: all 25 contested winners since 2022 recorded a valid apogee and recovered.

Weaker evidence. Several patterns are less secure. Individual school trends can’t be separated from noise (6 improving and 5 declining schools against about 2 in each direction expected by chance). The link between a school’s flight result and its next class change is not significant once the four tests are adjusted for each other. Among the regions, only Canada’s fall is clear. The CubeSat comparison and the link between design and build and flights are associations, and design and build is judged at the event, so neither can be read as the cause of a better standing or a better flight. The rubric changes (2016, 2023, and 2025) and the limited detail in 2018 and 2021 affect every comparison across years.

Methodology

The results for each competition arrived as a separate file with a different layout, so most of the work was getting them into one table that can be compared across years. Nothing was dropped silently, and every judgement call is listed in the cleaning log in the appendix.

Data sources

Table 11 lists the ten files, how each is structured, and what had to be handled for each.

Table 11: The ten source files.

Year	File	Structure	Rows	What had to be handled
2016	2016.xlsx	2 sheets (Basic, Advanced)	52	Different seven component rubric, Basic and Advanced classes (no altitude or motor classes). “n/a” scores are 0 points.
2017	2017.xlsx	7 sheets (all teams and per class)	82	First year with 10k and 30k classes by COTS, SRAD solid, and SRAD hybrid/liquid. Penalties stored negative.
2018	2018.xlsx	1 sheet	99	Total score plus target, predicted, and actual apogee only, with no component breakdown. 1 demonstration entry.
2019	2019.xlsx	2 sheets (1 hidden)	107	Full component detail, bonuses, apogee, recovery flag. The hidden sheet holds registration and fee data (not used).
2021	2021.xlsx	2 sheets	76	Virtual conference scoring (poster video, design competency and decisions), no flights or apogees. Footer rows removed.
2022	2022.xlsx	4 sheets (same table, 4 sort orders)	93	Formula-free. Prediction placeholder 0.1 for no-fly. The 4 sheets are cell-for-cell identical.
2023	2023.xlsx	4 sheets (same table)	119	Placeholders 999999 and 9999999. 8 school names with encoding damage. One mis-keyed bonus total.
2024	2024.xlsx	4 sheets (same table)	122	Free text flight status. One team classed differently across sheets. Launch dates. Placeholder 99999 for no prediction. 1 demonstration entry.
2025	2025.xlsx	1 sheet	143	Formulas with empty cached values (non-flyers), numbers stored as text, late admin penalty netted into the report score, and hard-coded rank columns that went stale. Multi-stage and 45k classes appear.
2026	2026.pdf	5 page PDF table	141	Parsed from the PDF text; every row reconciles. No team names or predictions. Penalties negative. 1 demonstration entry.

Organizer results

The overall winners and category winners declared by the organizers were transcribed from the Experimental Sounding Rocket Association’s past-events page (see Organizer sources) and matched to the score files by year and team number. They are used for the overall winner and category winner results only, and every other result uses the score files. The page gives a declared overall winner for each competition from 2017 and a first place for 55 category-years, and for 2016 it lists Basic and Advanced winners without team numbers. How the organizers’ lists compare with the highest scores is described in the category winners subsection.

Cleaning

These are the cleaning problems that mattered most.

1. **The rubric changed every few years.** 2016 used a seven component Basic and Advanced

scheme, 2017 introduced the 10k and 30k classes by COTS and SRAD motor type, 2021 was a virtual competition with no flights, 2023 split build quality from design quality, and 2025 added the multi-stage and 45k classes. Components were mapped to common buckets (admin, report, design and build, flight, bonus, penalty) and the scores were put on a common scale, as described below.

2. **Missing data was encoded as numbers.** “No data” appears as 999999 or 9999999 (2023), 99999 (2024), 0.1 (2022), and as an apogee of 0 or 1 ft. These would distort any average, so they were converted to missing values, and the flight analyses use only valid apogees.
3. **The workbooks repeat themselves.** The 2022 to 2024 files contain the same table four times in different sort orders. I checked that they match cell for cell (apart from rank tie-breaks) and used one copy.
4. **Stale and unreliable cells.** The 2025 rank columns are pasted values that no longer match the live Final Score formula for 83 of 143 teams (the top five are identical), some formula cells have no cached value, some numbers are stored as text, and one 2023 bonus total was mis-keyed. I recomputed ranks and bonuses from the components.
5. **Encoding damage.** 8 school names in the 2023 file are UTF-8 text that was read as Latin-1 (for example “KoÅ§ University” for Koç University), and were repaired.
6. **Identifying the same school.** Team numbers change every year, so the longitudinal unit is the school. 468 raw spellings were merged into 289 institutions using a reviewed alias table that covers typos, “The”, commas against dashes, and renamings such as Ryerson to Toronto Metropolitan University. Some years have schools with more than one team (2016 to 2018 and 2023 to 2026), so a school’s best entry in a year is used for trajectories.
7. **Penalty conventions.** Penalties are negative in 2017 and 2026 and positive in 2024 and 2025, and in 2025 the late admin penalty is netted inside the report score, which makes some report totals negative. All were restated as positive penalty points, with the report score shown gross.
8. **The 2026 PDF.** The results were extracted as text and parsed with a regular expression. All 140 rows are accounted for, the ranks are contiguous, and each row’s components sum to its total.

Making the years comparable

Raw totals can’t be compared across years. The winning total was 863 in 2017 and 1,118 in 2023, mostly because the bonuses changed. Three measures are used to deal with this. The *base score* is (total – bonuses) divided by that year’s base maximum, so bonuses don’t get counted as skill. *Component scores* are expressed as a share of each component’s cap. *Standing* is the percentile within the year’s field, which is not affected by any rubric change. Caps were read from the sheet headers where given and otherwise inferred from the highest observed scores, as shown in table 12.

Table 12: Point caps used for each year.

Year	Base max	Admin	Report	Design/build	Flight	Note
2016	1,400 (Adv. 1,430)	350	175	375	500 (Adv. 530)	Different rubric; buckets approximate
2017	1,000	100	200	200	500	
2018	1,000 (assumed)	–	–	–	–	Totals only; bonuses unknown, so the score is slightly overstated
2019	1,000 (plus up to 150 bonus)	60	200	240	500	
2021	1,000 (inferred)	100	400 (inferred)	500 (inferred, conference)	–	Virtual; not comparable to flight years
2022–2026	1,000 (plus bonuses)	60	200	240	500 (350 altitude, 150 recovery)	Bonuses: CubeSat, sportsman, prompt launch

The 2016 buckets combine several components (admin is deliverables, pre-flight, and safety, and design is presentation and technical application), so 2016 component comparisons are approximate and are drawn with hollow markers. The 2018 file reports only totals.

Definitions

- **Valid apogee** is a recorded apogee above 1 ft that isn't a placeholder value. **Recovered** means recovery points were awarded. The recovery points are always either 0 or 150, which is more reliable than the free text damage labels.
- **Altitude miss** is the absolute difference between apogee and target divided by the target, using the class's target altitude. The organizers' 2022 and later altitude score is $350 \times (1 - \text{miss}/30\%)$, floored at zero. My recomputation matches the sheets to within 0.004 points on 472 flights. **Prediction error** is the difference between apogee and the team's predicted apogee, divided by the prediction.
- **Class** is the target altitude combined with propulsion type (COTS, SRAD solid, SRAD hybrid/liquid, or multi-stage). SRAD hybrid/liquid is pooled across altitudes because of the small counts. A **switch** is a school's best entry at consecutive competitions being in a different class, where a **step up** is a higher altitude or a move from COTS to SRAD, a **step down** is the reverse, and **mixed** is one of each.
- **Paperwork and design score** is admin plus report plus design (plus the 2021 conference score), as a share of its cap.
- **Declared overall winner** is the team the organizers named overall winner. **Category winner** is the organizers' first place in an official category (target altitude by propulsion type, or Basic and Advanced in 2016), or the highest total score in the category where no first place is listed. A category-year is **contested** if it has at least 5 entries.
- **CubeSat bonus** is the 50 point bonus for a functional payload that adopts the CubeSat form factor (the payload bonus column in 2019 and 2022–25, and a bonus of 50 in the 2026

PDF). **Base standing** is the percentile of an entry’s base score within its competition (1 is the best), which excludes all bonuses.

- **Design and build** is the 240 point score judged at the competition from the poster session, discussions with the team, inspection of the hardware, and field preparation. The **technical report** (200 points) is submitted beforehand, although its analysis score can be adjusted after the officials’ interactions with the team.
- **Region** is the US, Canada, or Turkey, or otherwise Latin America (Brazil, Mexico, Argentina, Colombia, Paraguay), Europe (Poland, United Kingdom, Switzerland, Greece, Italy, Romania, Germany, Norway, Denmark, Spain), Oceania (Australia, New Zealand), or Asia, the Middle East and Africa (India, Egypt, South Korea, Thailand, Malaysia, Algeria, Taiwan, United Arab Emirates, Philippines, Nepal). One 2021 entry with no known country is counted in the last group. Turkey is kept separate because it is the third largest national contingent.

Statistical methods

Entries from the same school aren’t independent, so confidence intervals on trends and medians come from resampling schools (a cluster bootstrap), and intervals on proportions are Wilson intervals. Group comparisons use permutation tests, Fisher exact tests, or Mann–Whitney tests as noted in the text, and rank correlations are Spearman. Trends are linear slopes per year. With 7 to 10 data points and rubric changes along the way, they describe direction and rough size and should not be used as forecasts. The comparisons between eras pool 2017 and 2019 against 2022 to 2026. Regression to the mean in the switching analysis was handled by fitting the relationship between a school’s standing at one appearance and the next on schools that stayed in their class, then subtracting what that relationship predicts. The random seed for the resampling is fixed. The four class-switching tests are adjusted for each other with Holm’s method, the six regional comparisons are checked with a Bonferroni adjustment, and the other tests are not adjusted, so borderline results (p of about 0.03 to 0.05) are described as suggestive. Rank correlations between paperwork scores and flight results are computed on percentile ranks within each competition and pooled across competitions, and the area under the ROC curve is the probability that a randomly chosen entry with a valid apogee has a higher score than a randomly chosen entry without one. 2018 and 2021 are excluded from the absolute score slopes because 2018 has no bonus detail and 2021 had no flights.

Validation

935 of 935 rows reconcile, meaning admin plus report plus design plus conference plus flight plus bonus minus penalty equals the reported total, with a maximum residual of 0.03 points (2018 is excluded because it has no breakdown). The altitude score formula reproduces the organizers’ numbers on 472 flights. The recomputed top five lists match the sheets in every year from 2022 to 2026. For 2018, which only has totals, the total score correlates with having flown (Spearman $\rho = 0.66$), as it does in every other year (0.55 to 0.74), so that file doesn’t appear to be scrambled. The organizers’ listed category winners match the highest score in 50 of 55 category-years, and the five exceptions are explained in the category winners subsection.

Limitations

- **Entries aren’t attendance.** The files list scored entries, which sometimes include teams that never launched. 2016 has only 52 rows, so the earlier years may be incomplete. Partici-

pation trends reflect the files and not necessarily registrations.

- **Rubrics changed.** Putting scores on a common percentage scale removes the obvious distortions but not differences in what was being judged, how strictly, or by whom. Component trends, especially design and build, may partly reflect changes in judging.
- **Left-censoring.** A school first seen in 2016 or 2017 isn't necessarily new, so the experience analysis is restricted to schools first seen in 2018 or later, and the appearance counts are a lower bound on a school's history.
- **A school isn't a team.** Student teams, leadership, and designs turn over, so a school's trajectory mixes learning, turnover, and luck. Schools with more than one team are represented by their best entry.
- **Small cells.** Hybrid/liquid, 30k SRAD solid, and multi-stage have between 5 and 17 entries a year. Their year by year values are only shown for $n \geq 5$ and their trends have wide intervals.
- **Descriptive, not causal.** The relationship between flight success and class changes, the cost of stepping up, the link between design and build scores and flight outcomes, and the CubeSat comparison are associations. Weather, rule changes, and team composition aren't observed. Several results are borderline, and only the class-switching tests are adjusted for running several tests.
- **Winners.** The organizers' page gives no basis for the overall winner in 2023 and 2025, when the declared winner was not the highest scorer, and its 2023 10k hybrid/liquid line was not usable, so that category-year uses the highest score.
- **2026.** The PDF has no team names, predictions, or per-penalty detail, so the prediction analyses end in 2025.
- **Design and build is judged at the event.** The judges see the hardware shortly before it flies, so the link between design and build scores and flights partly reflects readiness to fly and is not evidence that design quality causes flight success.
- **CubeSat bonus.** The 2026 PDF has one bonus column, so a bonus of 50 is read as the CubeSat bonus, and the 2018 file has no bonus detail.

Organizer sources

- Experimental Sounding Rocket Association, Past IREC Events (winners, official documents): <https://www.esrarocket.org/past-irec-events>, read 16 October 2026.
- IREC 2027 Rules and Requirements, version 1.0 (effective 12 October 2026): section 3.3.5 (two teams from one institution), section 6.7.4 (payload form-factor bonus), section 6.8 (overall award), and appendix B.2 (design and build rubric). <https://static1.squarespace.com/static/687d3841f7d71450d1ba824d/t/6acd7cdb1c81c37d667ccfbd/1791851739101/IREC+2027+Rules+and+Requirements+v.1.0.pdf>
- Spaceport America Cup IREC Rules and Requirements Document, Rev. K draft (14 April 2022), including its revision history (the May 2018 change on multiple teams and the CubeSat bonus), the one team per university statement, the overall winner award, and the design and build scoring.

- ESRA, “Exciting Updates for IREC and ESRA’s Future Challenges!” (18 October 2024), on two teams from one university for the 2025 competition: <https://www.herox.com/IREC2025/update/7258>

Appendix

Cleaning log

Table 13: Cleaning and QA log.

Year	Issue	Action	Rows
2023	UTF-8 text decoded as Latin-1 in school names (e.g. ‘KoÅ§ University’)	repaired by cp1252→utf-8 round trip	8
all	Same institution spelled differently across/within years (case, ‘The’, hyphen vs comma, typos, renamed schools e.g. Ryerson → Toronto Metropolitan)	merged 181 name variants into 80 canonical institutions via reviewed alias table	413
2018, 2024, 2026	Non-competing demonstration entries mixed into the scoring sheets	flagged flag_demo; excluded from competitive analyses	3
2024	Oregon State (team 68) class differs between sheets of the same workbook	used the class shown on the other 3 sheets (SRAD Hybrid)	1
2018, 2019	Explicit target altitude contradicts class label (2018 Stanford grad class: 30k target / ‘10k’ class; 2019 Laval: 10k target / ‘30k’ class, flew 26,300 ft)	trusted class label	2
2023	Actual altitude 999999 (footer: ‘corrupt/no data’)	set to NaN	1
2018, 2019, 2022, 2024, 2025	Actual apogee 0 or 1 ft used as ‘no valid apogee recorded’ (1 ft only in 2025)	set to NaN; flagged via flew=False	112
2019, 2022, 2024	Predicted apogee 0 / 0.1 / 99999 = ‘no prediction submitted’	set to NaN	44
2022-2026	Free-text flight-status labels with many spellings (‘Excessive Damagae’, ‘no Damage’, ‘None’ meaning no damage...)	mapped to damage_class; ‘recovered’ derived from recovery points (strictly 0/150) rather than text	549
2023	ITU team 43 ‘Bonus Total’ cell (150) disagrees with its components (165) while Final Score already includes 165	bonus_pts recomputed from components	1
2017, 2026 vs others	Penalty sign convention differs (negative in 2017 & 2026, positive in 2024/25)	all penalties stored as positive magnitudes	
2025	Late-admin penalty is netted inside the ‘Total Project Technical Report’ column (can go negative, e.g. -43)	report_pts restated gross (Completeness+Style+Analysis); penalty moved to penalty_pts	110
2016	‘n/a’ scores	treated as 0 points (verified: reported totals only reconcile with n/a=0)	
2016	Different 6-component rubric (Basic/Advanced; no altitude classes)	components collapsed into 4 harmonised buckets - bucket comparisons with later years are approximate	52
2018	Only a total score is given (no component breakdown)	component columns left blank	99
2021	Virtual/no-flight scoring (poster video, design competency/decisions); no apogees	kept; conf_pts column; excluded from flight analyses	76

Year	Issue	Action	Rows
2022, 2023, 2024	Each workbook repeats the same table 4 times (Overall / By Class / Barrowman / By Team)	verified equal cell-by-cell (only rank tie-breaks differ); used one sheet	
2021, 2023, 2024	Footer rows (averages, sentinel legends) inside data range	excluded	
2025	Numbers stored as text ('0'); 91 formula cells with empty cached value (non-flyers)	coerced to numeric / blank=no flight	
2026	Source is a PDF (text extraction, regex row parse)	validated: 141/141 ranks contiguous and every row's components reconcile to its total	141
2018	No bonus/penalty breakdown, so base score = total/1000 (overstates teams that earned bonuses)	score_pct flagged approximate	99
all	Rubrics and bonus rules change almost every year (max total 862 in 2017 vs 1118 in 2023)	added score_pct = (total - bonuses)/base maximum and per-component % of cap (caps inferred from max observed; 2021 report/conference caps 400/500 are inferred)	
2022	Organiser rank column disagrees with rank recomputed from total score (tie-breaks / small offsets)	used recomputed ranks; top-5 identical to sheet: True	7
2023	Organiser rank column disagrees with rank recomputed from total score (tie-breaks / small offsets)	used recomputed ranks; top-5 identical to sheet: True	10
2025	Organiser rank column disagrees with rank recomputed from total score (rank columns are hard-coded values (not formulas) in the workbook and went stale after scores changed; Final Score is a live formula)	used recomputed ranks; top-5 identical to sheet: True	83
2017–2026	Organizer winners (past-events page) differ from the highest total score in some years and categories	overall winner reported as declared, with its rank by score; category winner uses the listed first place, and 1 unusable listing uses the highest score	5

Data dictionary

The cleaned table has one row per team entry, with the columns in table 14.

Table 14: Columns of the cleaned dataset.

Column(s)	Meaning
year, team_id	Competition year; per-year team number (not stable across years, so schools are the longitudinal unit).
school, school_raw, country	Canonical institution (alias-merged), the string as it appeared, and country.

Column(s)	Meaning
category, alt_class, propulsion	Harmonised class: target altitude (10k, 30k, 45k) by propulsion (COTS, SRAD solid, SRAD hybrid/liquid, multi-stage).
total_pts, score_pct	Reported total; (total minus bonuses) divided by the year's base maximum.
admin_pts ... penalty_pts (and *_pct)	Harmonised point buckets (penalties as positive magnitudes) and their share of cap.
target_alt_ft, pred_ft, apogee_ft	Target, predicted, and recorded apogee (placeholders removed).
flew, recovered, damage_class	1 if a valid apogee was recorded; 1 if recovery points were awarded; harmonised damage label.
alt_err_abs / signed, pred_err_abs / signed	(apogee - target) / target and (apogee - prediction) / prediction.
rank_in_year, rank_in_category, pct_rank	Ranks recomputed from total score; pct_rank is the 0 to 1 percentile within the year (1 is the winner).
competitive, flag_demo, no_deliverables	Analysis scope flags (demonstration entries are excluded).
prior_inperson_years, school_first_year_seen	School history within this dataset (a lower bound for activity before 2016).

AI Usage Statement

This document was compiled with the help of artificial intelligence (AI) tools, which were used for data processing, analysis, figure generation, and drafting of the text.