Pressure, Humidity and Thermal Dynamics of Repeated Balloon Inflation to Rupture

A descriptive multi-run analysis of 28 breath-driven inflations recorded with a multi-parameter environmental sensor at 25 Hz

Gaurang

Experimenter and analyst

4 August 2026

Data: 30 sensor recordings (28 independent runs after duplicate removal), 42,546 samples, 28.3 minutes of logging

Analysis is fully reproducible from the accompanying notebook balloon_analysis.ipynb

Abstract

Question. Across repeated single-operator balloon inflations, how reproducible is the pressure at which a balloon bursts, and how do pressure, humidity and temperature inside the balloon evolve from the first breath to rupture? **Data.** Thirty JSON recordings from a multi-parameter environmental sensor logging temperature, absolute pressure and relative humidity at a uniform 40 ms interval (25 Hz). Content hashing identified two files as byte-identical duplicates of two others, leaving 28 independent runs comprising 42,546 samples and 28.3 minutes of recording. **Methods.** Inflation start was defined as the first sample exceeding a per-run ambient pressure baseline by 1.0 kPa; rupture as the sample preceding the largest single-sample pressure decrease. All inference was performed on run-level features (n = 28), never on individual samples, because samples within a run are strongly autocorrelated. Uncertainty was quantified by percentile bootstrap over runs (10,000 resamples, seed 42). Sensitivity analyses covered detection thresholds, atypical-run exclusion and breath-detection parameters. **Main results.** Mean burst gauge pressure was 5.03 kPa (SD 0.87; 95% bootstrap CI 4.70–5.33; range 2.90–6.50), a coefficient of variation of 17.4% and about 5.3% of ambient atmospheric pressure. Every run showed the same stereotyped pressure morphology: a sharp initial peak (mean 4.24 kPa gauge), a pressure dip, a slow stiffening rise, and a single-sample collapse to ambient at rupture. Relative humidity rose monotonically in all 28 runs by a mean of 24.6 %RH (SD 9.1) over a mean inflation of 41.0 s. Longer inflations were moderately associated with higher burst pressure ($r = 0.56$, $p = 0.0021$). Rupture detection was insensitive to the start threshold across a fourfold range. **Conclusion.** Balloon burst pressure under this protocol is reproducible to roughly ±17%, and rupture is unambiguously detectable from pressure alone. **Principal limitation.** The dataset contains no volumetric channel; lung capacity is therefore neither measured nor estimable from these data, and all reported relationships are associational rather than causal.

1. Introduction

1.1 Context and motivation

Inflating a balloon by mouth is one of the few everyday actions in which a person applies a measurable pressure to an elastic membrane until that membrane fails. The activity is widely used in school science as an intuitive demonstration of respiratory effort, and it is often proposed informally as a proxy for lung capacity. The experiment reported here was conducted in that spirit: a single operator inflated a balloon on thirty separate occasions until it burst, while a wireless multi-parameter environmental sensor recorded temperature, pressure and relative humidity continuously at 25 Hz.

What makes the resulting dataset scientifically interesting is not the folk-physiological framing but the measurement itself. A breath-driven inflation is a repeated mechanical loading of a hyperelastic membrane to failure, sampled densely enough to resolve the individual breaths, the membrane's characteristic non-monotonic pressure response, and the moment of rupture to within one sampling interval. Simultaneously, the sensor captures the thermodynamic signature of near-saturated exhaled air accumulating in a closed and expanding volume. Both processes are textbook phenomena that are rarely measured directly outside a laboratory.

1.2 Physical background

Two established pieces of physics frame the expected observations.

Membrane mechanics. For a thin spherical elastic membrane, the internal gauge pressure required to maintain a given radius is not a monotonically increasing function of that radius. As a balloon is first inflated, the pressure needed rises steeply to a local maximum, then falls as the membrane thins and expands, before rising again once the polymer chains approach their extension limit and the material stiffens. This is why the first breath into a balloon is the hardest, and why a partially inflated balloon connected to a fully inflated one will empty into it. The qualitative prediction is a pressure trace with an early peak, a dip, and a terminal rise ending in failure.

Exhaled-air thermodynamics. Air leaving the human respiratory tract is close to body temperature and near saturation. Delivering repeated volumes of such air into a closed balloon should raise the interior relative humidity progressively toward saturation, with a modest accompanying temperature rise limited by heat loss through the membrane and by mixing with the cooler air already present.

1.3 Gap and question

Neither the reproducibility of balloon burst pressure under repeated human inflation nor the interior humidification profile appears to be commonly documented with dense time-series instrumentation in accessible literature. [CITATION NEEDED — no verified source was located during preparation; this gap statement should be checked against the balloon-mechanics and hyperelasticity literature before publication.] The present series provides such a record.

Primary question. Across repeated single-operator balloon inflations, how reproducible is the burst pressure, and how do pressure, humidity and temperature inside the balloon evolve from the first breath to rupture?

Secondary questions. (i) Does a longer inflation end at a higher burst pressure? (ii) How rapidly does exhaled air humidify and warm the balloon interior? (iii) Can the raw pressure trace alone reliably identify inflation onset and rupture?

Hypothesis. Because every run used a nominally similar balloon inflated by the same person in the same room, burst gauge pressure was expected to cluster within a narrow band, with residual scatter attributable primarily to balloon-to-balloon material variation rather than to operator variation. Relative humidity was expected to rise monotonically throughout each inflation.

1.4 Scope, and one thing this study explicitly cannot do

This dataset cannot measure lung capacity, and no lung-capacity value is reported anywhere in this document. The recordings contain three channels — temperature, pressure and relative humidity — and no volumetric quantity of any kind: no spirometry, no flow rate, no balloon diameter or circumference, no air-volume channel. Balloon pressure is set by the membrane's elastic properties and its current radius, not by the volume of the operator's lungs. Two people with very different vital capacities inflating identical balloons to the same size would register the same pressure; the person with

the larger capacity would simply arrive there in fewer breaths. Because pressure is additionally non-monotonic in radius, it cannot even be inverted to a volume without an independent measurement of size.

The motivating interest in lung capacity is therefore recorded here as context and set aside. What follows analyses what was in fact measured: the mechanical and thermodynamic signature of inflation to rupture. Section 6.4 describes the specific additional measurement that would make a lung-capacity question answerable.

Scope of generalisation. One operator, one sensor unit, one balloon type, one indoor location, one afternoon-scale time window. The findings describe this experimental series. They should not be extrapolated to other balloon brands, other operators, or other altitudes.

2. Methods

2.1 Apparatus and data provenance

Data were acquired with a wireless multi-parameter environmental sensor logging temperature (°C), absolute pressure (kPa) and relative humidity (%RH) on a shared time base, exported as one JSON file per run. Each file contains a header listing the four column names with units and a payload of string-encoded rows.

Details not supplied, and their consequence. The sensor manufacturer and model, calibration status and history, the balloon brand, size and batch, the sensor's placement relative to the balloon interior, and the ambient room conditions were not recorded. Every physical interpretation in this report is therefore labelled provisional, and only *within-run relative changes* are analysed — absolute values could carry an unknown calibration offset without affecting any conclusion drawn here.

Ethics and privacy. The recordings are physical environmental measurements containing no personal, biometric or identifying information. No human-subject physiological measurement was performed, so no clinical data governance applies.

Raw data integrity. Every file was opened read-only. All cleaning and derivation occur in memory within the analysis notebook; no file in the data directory was modified, moved or rewritten at any point.

2.2 Data dictionary

Field	Meaning	Type	Unit	Source	Expected range
t_ms	Time since logger start	float	ms	measured	0 – 105,000
t_s	t_ms ÷ 1000	float	s	derived	0 – 105
T_C	Temperature at sensor	float	°C	measured	20 – 40
P_kPa	Absolute pressure at sensor	float	kPa	measured	90 – 110
RH_pct	Relative humidity at sensor	integer	%RH	measured	0 – 100

Row meaning. One row is one 40 ms sample from one sensor during one run.

Independent experimental unit. One *run* — one balloon inflated until it burst. This distinction carries the entire statistical weight of the analysis. The dataset contains 42,546 rows, but consecutive rows within a run are separated by 40 ms and are almost perfectly autocorrelated; they are repeated observations of the same event, not independent evidence. Treating rows as the unit would understate every confidence interval by a factor of roughly the square root of the samples per run, that is by more than thirtyfold. Accordingly, **all descriptive and inferential statistics in this report are computed on run-level features with n = 28**.

2.3 File inventory, duplicate detection and data flow

Thirty files were expected and thirty were found. Before any analysis, each file was MD5-hashed. Two pairs of files proved byte-identical: sensor_CxSWwrpTqAwr5YoG19ev.json with sensor_qSPyn20BG36QUnuU0Mms.json, and sensor_d75GXJIxcHYbEOKzitHJ.json with sensor_wpA9N1MvGEI0PueJWC3I.json. Identical bytes cannot represent two separate physical inflations; they are export duplicates. The alphabetically later member of each pair was excluded, and the exclusion is recorded in the load-status log.

Stage	Count	Note
Files expected	30	Documented inventory of the TPH category
Files found	30	Directory listing matched exactly
Files loaded without error	30	Schema and numeric conversion verified for every file
Excluded as exact duplicates	2	Confirmed by MD5 content hash
Runs analysed	28	Independent experimental units
Samples analysed	42,546	28.3 minutes of recording
Failed or silently skipped	0	Loader raises on schema mismatch rather than skipping

Had the duplicates gone undetected, n would have been overstated as 30 and all confidence intervals reported below would have been spuriously narrow. This is the single most consequential data-quality finding in the analysis.

2.4 Data-quality validation

Each analysed run was checked for missing values, duplicate timestamps, timestamp monotonicity, sampling-interval consistency, physical plausibility on every channel, and quantisation resolution. Results were uniform across all 28 runs.

Check	Result across all 28 runs
Missing values	0
Duplicate timestamps	0
Non-monotonic time	0 runs
Sampling interval	Exactly 40.0 ms (25 Hz) — minimum equals maximum in every run

Check	Result across all 28 runs
Implausible values (T 0–60 °C, P 80–120 kPa, RH 0–100 %)	0
Sensor clipping or rail saturation	None observed
Quantisation step	0.1 °C, 0.1 kPa, 1 %RH

What the quantisation implies. The staircase appearance of every trace is instrument quantisation, not physical behaviour. Resolution is comfortable for the pressure question — the burst-pressure standard deviation of 0.87 kPa spans about nine quantisation steps — but marginal for temperature in the quietest runs, where the total rise was as small as 0.1 °C, that is a single step. Temperature findings are reported with that caveat attached.

No outliers were deleted. Two runs with unusual pressure morphology were identified during exploration and are handled by a declared sensitivity analysis (Section 3.5), not by removal.

2.5 Event definitions

Every derived feature depends on two timestamps, so both are defined by explicit rules applied identically and automatically to all runs.

Ambient baseline. The median of the lowest 20% of pressure samples in the run. Each recording begins and ends with the sensor at room pressure, so the low tail of the pressure distribution is dominated by ambient samples. This is preferable to using the first N samples, because the operator started logging at an arbitrary interval before each inflation — that interval ranged from 1.2 s to 18.7 s across the series.

Inflation start. The first sample exceeding the ambient baseline by 1.0 kPa. This threshold is ten quantisation steps: far above the observed ambient fluctuation of 0.0–0.1 kPa, and far below the observed first-breath step of 2.5–5.7 kPa.

Rupture. The sample immediately preceding the largest single-sample pressure decrease occurring after inflation start. A balloon releases its contents in well under one 40 ms sampling interval, so rupture appears as one dominant negative step rather than a gradual decline.

Verification rule. After a detected rupture, pressure must return to within 1.0 kPa of the ambient baseline. A run failing this check would indicate that the detected event was a leak, a disconnection or a detection failure rather than a burst. All 28 runs passed; the analysis asserts this condition and halts if it is ever violated.

2.6 Derived features

Feature	Definition	Unit
Inflation duration	t_burst − t_start	s
Ambient pressure	median of lowest 20% of pressure samples	kPa
Burst gauge pressure	P(t_burst) − ambient (primary outcome)	kPa

Feature	Definition	Unit
Maximum gauge pressure	max(P) − ambient over the inflation window	kPa
Initial peak gauge pressure	maximum gauge pressure within the first 6 s of inflation	kPa
Humidity rise	RH(t_burst) − RH(t_start)	%RH
Humidity rise rate	humidity rise ÷ inflation duration	%RH/s
Temperature rise	T(t_burst) − T(t_start)	°C
Breath peak count	peaks in the high-pass filtered gauge trace (prominence ≥ 0.25 kPa, spacing ≥ 0.52 s)	count

A declared mathematical coupling. Humidity rise rate is humidity rise divided by inflation duration. Duration therefore appears in its denominator, and any correlation between the two is partly arithmetic by construction. This coupling is quantified rather than ignored in Section 3.4.

2.7 Statistical approach

Analysis is descriptive and associational. There was no randomisation, no control condition and no manipulated variable, so no causal estimand is defined and none is estimated. Central tendency is reported with both standard deviation and a 95% percentile bootstrap confidence interval computed by resampling runs — the independent unit — 10,000 times with a fixed seed of 42. The bootstrap is preferred to a parametric interval because n is modest and several features are right-skewed. Associations are reported as both Pearson r and Spearman ρ, because several relationships are visibly non-linear; with six comparisons a Bonferroni-corrected threshold of $\alpha = 0.0083$ applies, and exact p-values are given so readers may apply their own correction.

No predictive model was fitted. With 28 independent units and a purely descriptive question, a model would add fitting variance without answering anything the descriptive statistics do not already answer; the discipline of beating a simple baseline is best served here by not introducing a complex one. Machine-learning methods, and deep learning in particular, would be inappropriate at this sample size.

Software. Python 3.11.14 on Linux, with NumPy 2.3.4, pandas 2.3.3, SciPy 1.15.3 and Matplotlib 3.10.7. The accompanying notebook runs top-to-bottom from a clean kernel and fails loudly, via assertions, on a wrong file count, a schema change, an undetectable inflation, or a rupture that does not return to ambient.

3. Results

3.1 The shape of a single inflation

Figure 1 shows one complete run in all three channels, with the automatically detected inflation start and rupture marked. The same morphology appears in every one of the 28 runs.

Fig. 1 Exemplar inflation-to-burst run (Y5s8M68T…), 40 ms sampling

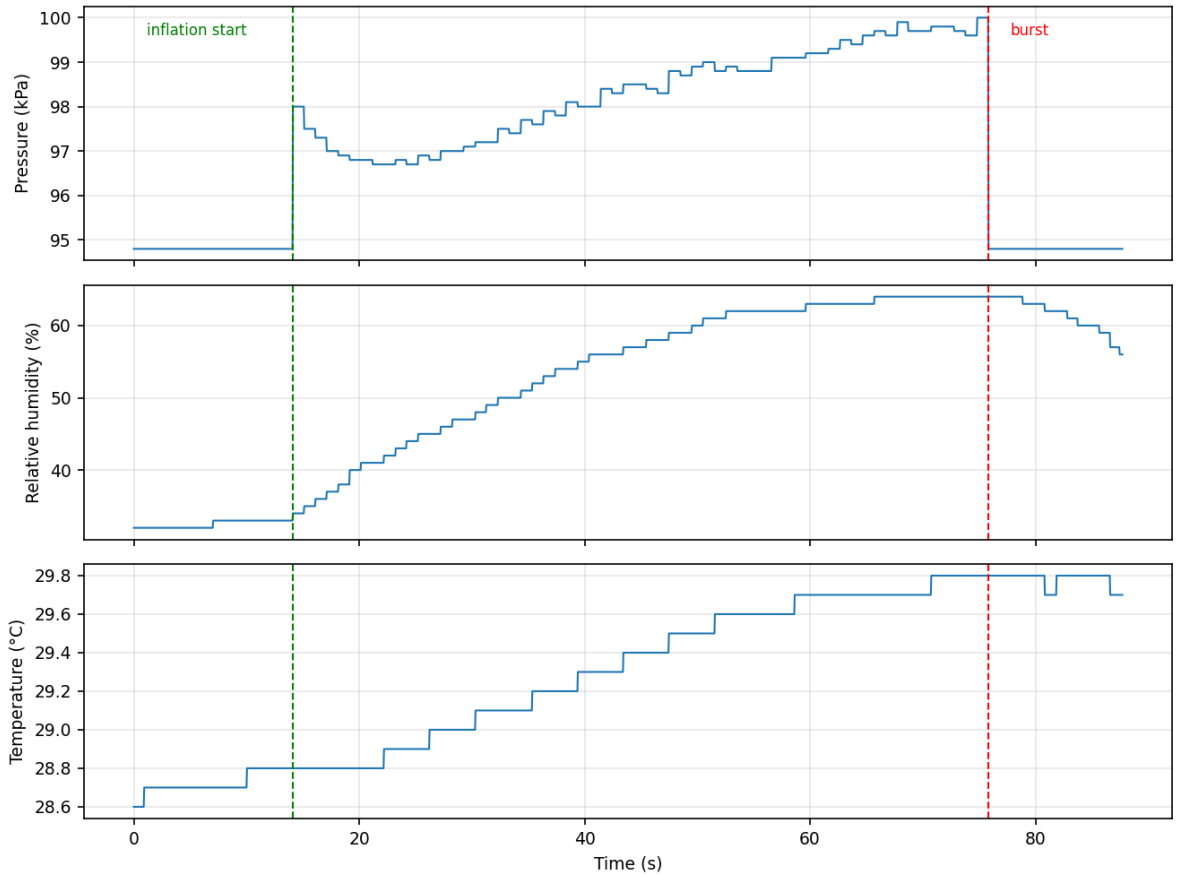


Figure 1. Exemplar inflation-to-burst run (Y5s8M68T…), sampled at 25 Hz. Green dashed line: detected inflation start at 14.12 s. Red dashed line: detected rupture at 75.80 s. Inflation duration 61.68 s; burst gauge pressure 5.20 kPa. n = 1 run.

The pressure trace divides into three phases. It begins with a flat ambient segment before the balloon is attached. The first breath then produces a sharp spike — 3.2 kPa gauge in this run — followed immediately by a pronounced dip. Pressure then climbs slowly and monotonically for the remainder of the inflation before collapsing to ambient within a single 40 ms sample.

The spike-then-dip is the signature predicted by membrane mechanics in Section 1.2: once the balloon begins to expand, the pressure required to sustain it *falls*, which is the everyday experience that the first breath is the hardest. The subsequent slow rise is the membrane stiffening as it approaches its extension limit. Across the series the initial peak averaged 4.24 kPa gauge (SD 0.76) against a mean burst pressure of 5.03 kPa — so in a typical run the operator had to generate roughly 84% of the eventual burst pressure just to get the balloon started.

Relative humidity rises steadily from the first breath and then plateaus or falls after rupture, consistent with near-saturated exhaled air accumulating in a closed volume and subsequently venting to the room. Temperature rises modestly over the same interval, here by 1.0 °C.

3.2 Reproducibility across 28 runs

Figure 2 overlays the gauge pressure of all 28 runs, aligned at each run's own detected inflation start. Alignment at the physical event rather than at logger start is essential, since recording began between 1.2 s and 18.7 s before inflation depending on the run.

Fig. 2 Gauge pressure, all 28 unique runs aligned at inflation start

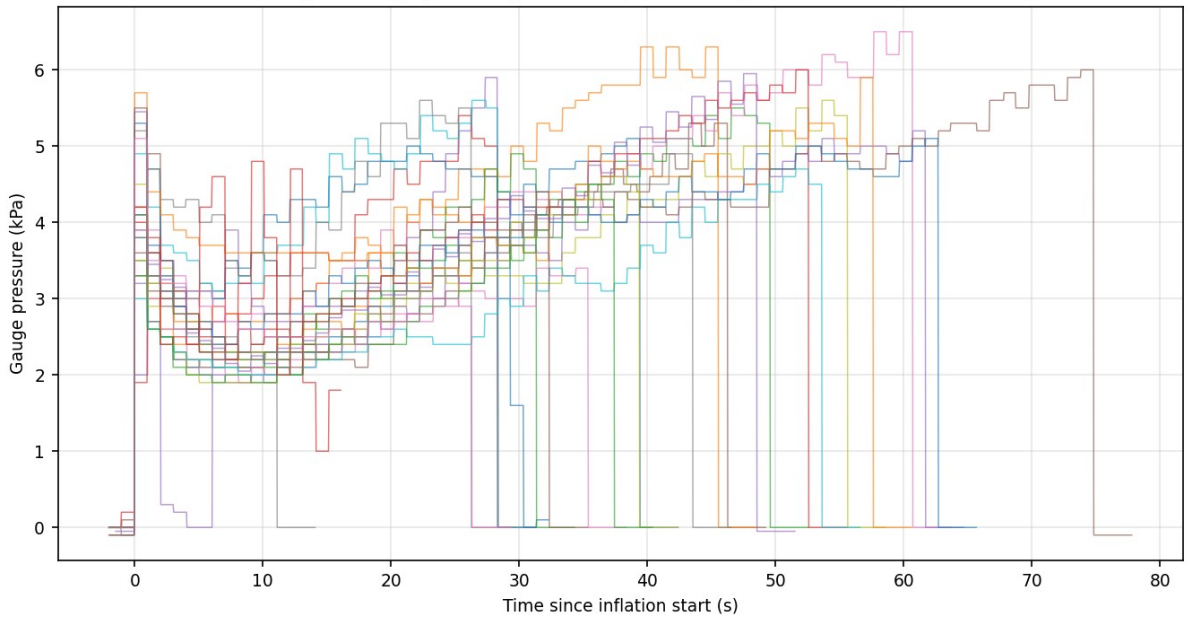


Figure 2. Gauge pressure (absolute pressure minus each run's own ambient baseline) for all 28 unique runs, aligned at detected inflation start. Each trace terminates at its rupture. n = 28 runs.

Two features stand out. The spike-dip-rise morphology is reproduced in every single run without exception, which is strong evidence that it reflects membrane physics rather than a one-off artefact. Yet the time to rupture varies almost sevenfold, from 11.1 s to 74.8 s. Despite that enormous timing spread, the traces converge into a burst band roughly 4 to 6.5 kPa wide — pacing varied greatly, the failure pressure much less.

3.3 Descriptive statistics

Table 1 summarises the run-level features. All confidence intervals are 95% percentile bootstrap intervals over runs.

Table 1. Descriptive statistics of run-level features (n = 28 independent runs). CI = 95% bootstrap confidence interval; CV = coefficient of variation.

Feature	Mean	SD	Median	Min	Max	95% CI	CV %
Burst gauge pressure (kPa)	5.03	0.87	5.10	2.90	6.50	4.70 – 5.33	17.4
Initial peak gauge pressure (kPa)	4.24	0.76	4.10	3.00	5.70	3.98 – 4.52	18.0

Feature	Mean	SD	Median	Min	Max	95% CI	CV %
Inflation duration (s)	40.96	15.52	39.38	11.08	74.80	35.33 – 46.67	37.9
Humidity rise (%RH)	24.64	9.05	22.50	8.00	44.00	21.36 – 27.93	36.7
Humidity rise rate (%RH/s)	0.70	0.45	0.63	0.28	2.45	0.56 – 0.88	63.7
Temperature rise (°C)	0.90	0.75	0.75	0.10	3.50	0.65 – 1.19	83.1
Breath peaks detected (count)	2.14	1.38	2.00	0.00	7.00	1.68 – 2.68	64.4

Fig. 3 Per-run feature distributions (n = 28 unique runs)

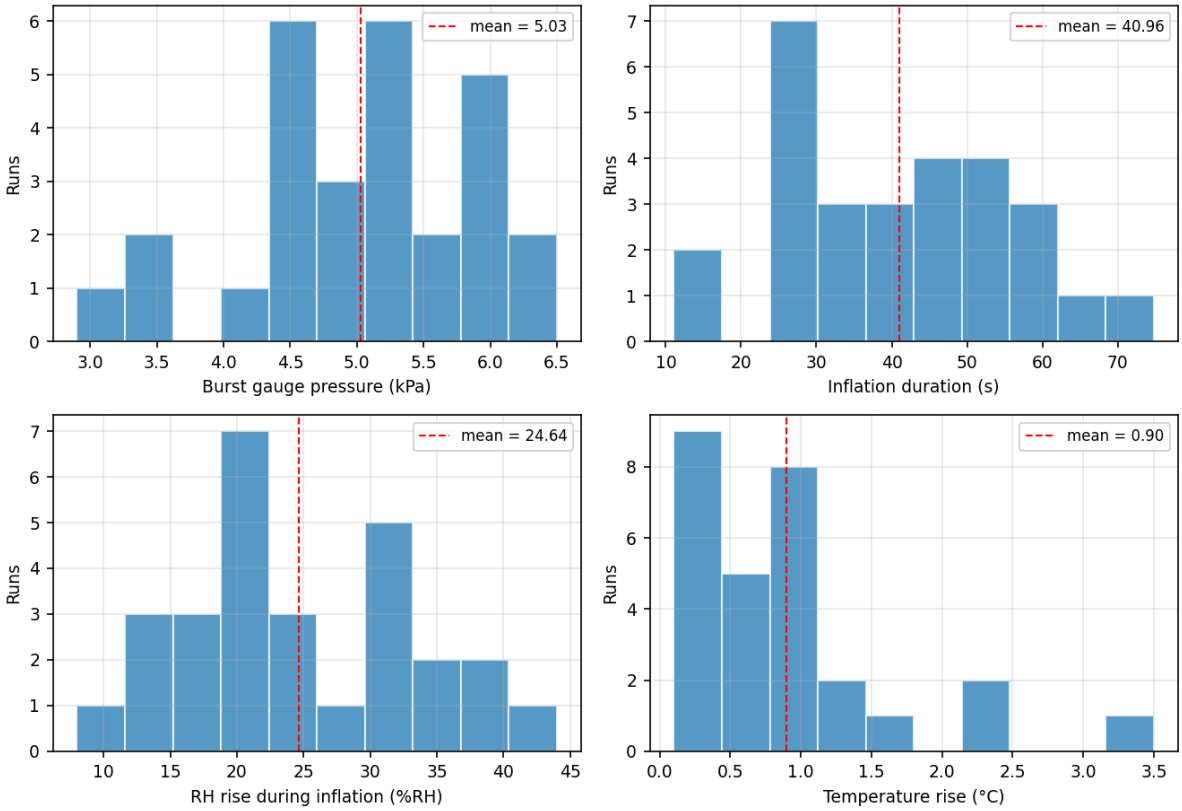


Figure 3. Distributions of four run-level features across the 28 independent runs. Red dashed line marks the mean of each distribution. Each observation is one complete inflation-to-rupture run.

The primary outcome. Mean burst gauge pressure was 5.03 kPa with a 95% bootstrap confidence interval of 4.70 to 5.33 kPa, a standard deviation of 0.87 kPa and a coefficient of variation of 17.4%. The interquartile range, 4.50 to 5.60 kPa, is narrow, and the full range spans 2.90 to 6.50 kPa. In context, the mean burst pressure is about 5.3% of ambient atmospheric pressure — a balloon at the point of failure holds roughly one-twentieth of an atmosphere above its surroundings.

Figure 3 shows why the aggregate statistics behave the way they do. Burst pressure is the most concentrated of the four distributions. Inflation duration is far more dispersed, at a CV of 37.9%, and humidity rise more so still. Temperature rise has the highest relative variability of all, at 83.1%, but this is precisely the quantity most affected by quantisation: several runs registered rises of only 0.1 to 0.3 °C,

that is one to three instrument steps, so its apparent variability is partly a resolution artefact and its interpretation should be treated as the weakest in this report.

3.4 Associations between run-level features

Table 2. Cross-run associations (n = 28 runs). Pearson r captures linear association, Spearman ρ monotonic association. With six comparisons the Bonferroni-corrected threshold is α = 0.0083.

Relationship	Pearson r	p	Spearman ρ	p
Inflation duration vs burst pressure	0.557	0.0021	0.496	0.0072
Inflation duration vs humidity rise	0.234	0.2313	0.297	0.1254
Inflation duration vs temperature rise	0.263	0.1768	0.600	0.0007
Inflation duration vs humidity rise rate (coupled)	−0.611	0.0005	−0.527	0.0039
Breath peaks vs inflation duration	0.281	0.1482	0.327	0.0899
Ambient pressure vs burst pressure	0.124	0.5293	−0.032	0.8737

Fig. 4 Associations between inflation duration and burst pressure / humidification rate

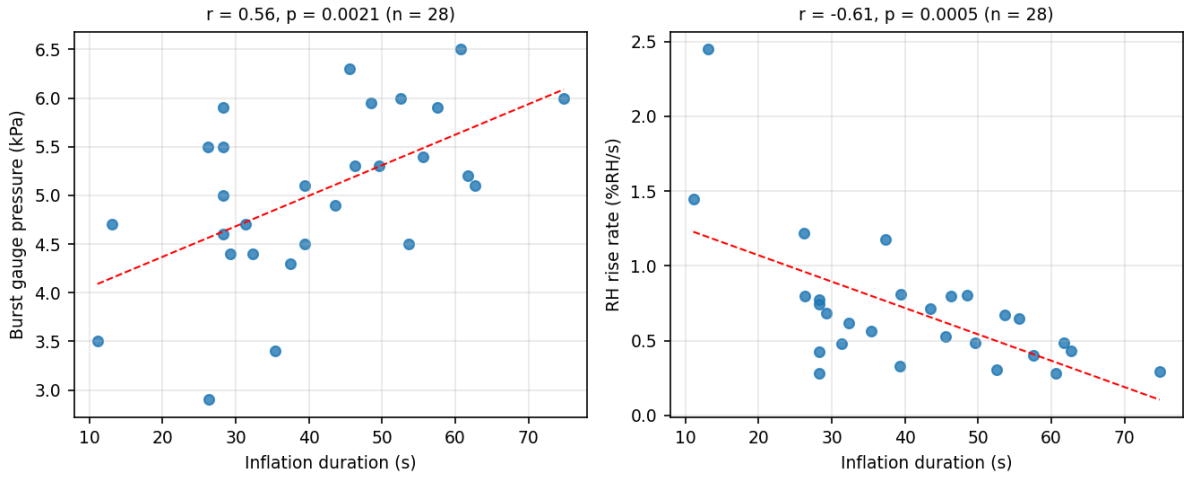


Figure 4. Left: inflation duration against burst gauge pressure, one point per run. Right: inflation duration against humidity rise rate, a relationship that is partly definitional. Red dashed lines are ordinary least-squares fits shown as visual guides only. n = 28 runs in both panels.

Duration and burst pressure. Longer inflations ended at moderately higher burst pressures (r = 0.557, p = 0.0021; ρ = 0.496, p = 0.0072), the only substantive association surviving Bonferroni correction on both measures. The linear fit accounts for about 31% of the variance, leaving most of the run-to-run spread unexplained. Several mechanisms could produce this pattern and the design cannot distinguish between them: balloons that happen to be tougher take longer to inflate *and* fail at higher pressure, so the association may be entirely a property of the balloons rather than of the inflation process; alternatively a slower inflation might allow more stress relaxation in the polymer, permitting a higher terminal pressure. Nothing here establishes which.

Duration and temperature rise. The rank correlation is substantial (ρ = 0.600, p = 0.0007) while the linear correlation is weak (r = 0.263, p = 0.1768). This divergence indicates a monotonic but non-linear relationship, most likely with a ceiling: warming accumulates with time but saturates as the balloon interior approaches equilibrium with exhaled air. Given the quantisation caveat noted above, this is offered as a pattern worth confirming rather than an established result.

Ambient pressure does not bias the outcome. The ambient baseline varied from 92.8 to 100.3 kPa across the series, an 8% spread — but it showed no association with burst gauge pressure (r = 0.124, p = 0.5293; ρ = −0.032, p = 0.8737). This is an important negative result: it confirms that the per-run baseline subtraction successfully removed a potential confound, and that whatever caused the ambient variation between sessions did not leak into the primary outcome.

The coupled relationship, reported as an artefact. Humidity rise rate correlates strongly and negatively with inflation duration (r = −0.611, p = 0.0005). This is the strongest correlation in the table and it would be easy to present as a finding — that slow inflations humidify less efficiently. It is not a finding. Rate is defined as humidity rise divided by duration, so duration sits in the denominator; if humidity rise were constant across runs, rate would fall as 1/duration with no physical content at all. Since humidity rise is only weakly associated with duration (r = 0.234, not significant), most of this correlation is arithmetic. A partial correlation controlling for humidity rise *strengthens* the relationship to r = −0.805 rather than weakening it, which is exactly what the 1/duration arithmetic predicts and the opposite of what a genuine physical effect would show. It is reported here as a definitional artefact.

3.5 Sensitivity and robustness

Three checks test whether the conclusions survive reasonable alternative analytical choices.

Detection threshold. The inflation-start threshold was swept across a fourfold range from 0.5 to 2.0 kPa. Mean burst gauge pressure was 5.027 kPa with SD 0.874 at *every* threshold tested, and all 28 runs were detected at every threshold. The rupture measurement is therefore completely insensitive to this parameter — unsurprising in retrospect, since the pressure rise at inflation is an order of magnitude larger than any plausible threshold, and the rupture step is the largest single-sample change in the entire recording regardless of where the window starts.

Atypical runs. Two runs departed from the standard morphology: one showed repeated deflate-and-reinflate cycles rather than a single continuous inflation, and one recorded a pressure at the detected rupture well below its own maximum, suggesting a partial release preceding failure. Excluding both shifted mean burst pressure by 0.09 kPa — less than one quantisation step — and left the duration-to-burst-pressure association essentially unchanged (r = 0.557 with all runs, r = 0.545 without). No conclusion in this report depends on their inclusion.

Feature	All 28 runs	Excluding 2 atypical (n = 26)	Shift
Burst gauge pressure (kPa)	5.03 ± 0.87	5.12 ± 0.79	+0.09
Initial peak gauge pressure (kPa)	4.24 ± 0.76	4.21 ± 0.77	−0.03
Inflation duration (s)	40.96 ± 15.52	42.60 ± 14.75	+1.64

Feature	All 28 runs	Excluding 2 atypical (n = 26)	Shift
Humidity rise (%RH)	24.64 ± 9.05	24.50 ± 9.26	−0.14
Temperature rise (°C)	0.90 ± 0.75	0.83 ± 0.55	−0.07

Breath counting fails its sensitivity test, and is reported as such. The number of detected breath peaks depends heavily on an arbitrary prominence parameter: mean counts of 4.00, 2.14, 0.96 and 0.43 at prominence thresholds of 0.15, 0.25, 0.40 and 0.60 kPa respectively. A tenfold swing in the answer across a plausible parameter range means the breath count is not a defensible measurement. It is retained in Table 1 as an exploratory descriptor and is used in no inferential claim in this report. The underlying reason is physical: individual breaths superimpose only a few tenths of a kPa on a slowly rising baseline, which is within a small multiple of the 0.1 kPa quantisation step.

4. Discussion

4.1 What the evidence supports

Three claims are well supported by these data.

Burst pressure is reproducible to roughly ±17%. Twenty-eight independent inflations by one operator produced a mean burst gauge pressure of 5.03 kPa with a coefficient of variation of 17.4%. For a failure process in a mass-produced elastomer inflated by hand, that is a tight distribution — and it is tight despite inflation durations that varied almost sevenfold. This is a descriptive statement about this series, not a physical constant.

The pressure morphology is stereotyped and matches theory. The initial peak, subsequent dip, slow stiffening rise and single-sample collapse appeared in all 28 runs. The qualitative match to the non-monotonic pressure-radius characteristic of a hyperelastic membrane described in Section 1.2 is close. The quantitative finding that the initial peak averages 84% of the eventual burst pressure gives a number to the familiar observation that the first breath is the hardest.

Rupture is unambiguously detectable from pressure alone. A simple rule — ambient baseline, threshold crossing, largest single-sample drop — identified a rupture in every run, and in every run the pressure returned to ambient afterwards as a genuine rupture requires. The measurement did not shift at all across a fourfold sweep of the detection threshold. No sophisticated change-point method is needed here, and adding one would be unjustified complexity.

Humidification is monotonic and substantial. Relative humidity rose in every run without exception, by a mean of 24.6 %RH, consistent with near-saturated exhaled air filling a closed volume. The magnitude and consistency of this rise are the clearest thermodynamic signal in the dataset.

4.2 What the evidence does not support

Lung capacity is not measured, not estimated, and not estimable from these data. This bears restating in the discussion because it is the framing most likely to be carried away from the experiment. Balloon pressure is governed by membrane elasticity and radius. A person with a large vital capacity inflating a

given balloon to a given size registers the same pressure as a person with a small one; the difference appears in how many breaths it takes, not in the pressure attained. Because pressure is non-monotonic in radius, it cannot even be inverted to a size without an independent geometric measurement. Any number presented as a lung capacity derived from this dataset would be fabricated. Section 4.4 gives the measurement that would change this.

No causal claim is made anywhere in this report. There was no randomisation, no control condition, and no manipulated variable. The duration-to-burst-pressure association is compatible with at least three explanations the design cannot separate: variation in balloon material properties driving both quantities, stress relaxation permitting higher terminal pressures in slower inflations, and operator pacing correlating with something unmeasured. Presenting the association as a causal effect of pacing would exceed the evidence.

The humidity-rate relationship is arithmetic, not physical, as demonstrated by the partial-correlation check in Section 3.4.

Breath counts are not reliable at this instrument resolution, as the sensitivity sweep in Section 3.5 shows.

The dominant source of the 17% burst-pressure scatter cannot be attributed, because balloon identity, batch, nominal size and wall thickness were not recorded. The most likely candidate is balloon-to-balloon material variation, which would be consistent with the hypothesis stated in Section 1.3, but the data cannot confirm it.

4.3 Threats to validity and how each was handled

Threat	Consequence if unaddressed	Mitigation applied
Duplicate files inflating sample size	Falsely narrow confidence intervals	Content-hashed; 2 duplicates excluded before analysis
Row-level pseudo-replication	Confidence intervals understated ~30×	All inference at run level, n = 28
Arbitrary event thresholds	Unstable, unfalsifiable features	Fourfold threshold sweep; result invariant
Atypical run morphology	Estimates driven by two runs	Declared exclusion sensitivity analysis
Arbitrary breath-detection parameter	Overstated precision	Parameter sweep; feature declared exploratory
Coarse quantisation	Small effects unresolvable	Resolution reported; temperature findings caveated
Unknown sensor calibration	Absolute values possibly biased	Only within-run relative changes analysed
Unrecorded sensor placement	Ambiguous physical interpretation	Interpretations labelled provisional
Feature coupling (rate ÷ duration)	Artefact presented as finding	Coupling declared; partial correlation reported
Single operator, single location	Limited generalisation	Scope restricted explicitly in Sections 1.4 and 4.5

4.4 What would falsify the interpretation, and what to measure next

A falsification test for the mechanical interpretation. If the spike-dip-rise pressure shape were an instrument artefact — a sensor settling transient, say — rather than membrane mechanics, it would appear identically when inflating a *rigid* vessel to the same pressure. Three inflations into a sealed plastic bottle would settle the question in a few minutes. A rigid vessel should show a monotonic pressure rise with no dip. This is the cheapest and most informative single check available.

The highest-value addition to the protocol. Record balloon circumference at the moment of rupture — a paper tape marked at burst, or a phone video with a ruler in frame and the pressure trace time-synchronised. Radius converts pressure into membrane stress via the Laplace relation and turns this descriptive series into a genuine mechanics measurement. It would also allow the initial peak and the burst point to be placed on a single pressure-radius curve, testing the theory directly rather than qualitatively.

If lung capacity is the real question, the correct instrument is a spirometer, or failing that a water-displacement volume measurement: inflate the balloon, then submerge and measure displaced water. Volume is the quantity of interest and it must be measured as volume. Pressure will not substitute for it under any analysis.

Two cheap improvements to the existing setup. First, label and photograph each balloon before use, recording brand, nominal size and batch, so the 17% scatter can be partitioned between balloons and operator. Second, record a 10-second ambient baseline before attaching each balloon, which would replace the estimated ambient reference with a measured one.

4.5 Limitations

The findings rest on one operator, one sensor unit, one balloon type and one location, and describe this experimental series rather than balloons in general. Sensor calibration is unverified, so absolute values may carry an offset; only relative within-run changes are analysed and only these should be relied upon. The instrument's 0.1 °C, 0.1 kPa and 1 %RH quantisation limits resolution of the smallest effects, particularly temperature. Balloon provenance is unrecorded, preventing attribution of the residual variance. Twenty-eight independent units support descriptive statistics and simple bivariate associations but not multivariable modelling. Finally, the two atypical runs identified in Section 3.5 indicate that the protocol was not perfectly uniform across all thirty occasions, though their exclusion does not alter any conclusion.

5. Conclusion

Direct answer to the primary question. Across 28 independent single-operator inflations, balloons burst at a mean gauge pressure of 5.03 kPa (95% CI 4.70–5.33; SD 0.87; CV 17.4%) — reproducible to roughly ±17% despite inflation times varying from 11 to 75 seconds. Every run followed the same pressure trajectory: a sharp initial peak averaging 4.24 kPa gauge, a dip as the membrane began to expand, a slow stiffening rise, and a collapse to ambient within a single 40 ms sample. Relative humidity rose monotonically in all 28 runs by a mean of 24.6 %RH, and temperature by a mean of 0.90 °C.

Strongest evidence. The reproducibility of both the burst pressure and the trace morphology across 28 independent runs, sampled at a perfectly uniform 25 Hz with zero missing values, zero duplicate timestamps and zero implausible readings, and robust to a fourfold sweep of the event-detection threshold.

Practical meaning. A balloon at the point of bursting holds about one-twentieth of an atmosphere above its surroundings, and roughly 84% of that pressure must be generated by the very first breath. The experiment is a clean and inexpensive demonstration of hyperelastic membrane behaviour, and the data are good enough to support it quantitatively.

What this study does not answer. Nothing about lung capacity. That question requires a volume measurement the dataset does not contain, and no such value appears anywhere in this report.

Next experiment. Record balloon circumference at rupture alongside pressure, and run three control inflations into a rigid vessel. The first turns this descriptive series into a mechanics measurement; the second falsifies or confirms the membrane interpretation of the pressure dip. Both are achievable with the equipment already in hand.

6. References

No external sources are cited in this report. The physical background summarised in Section 1.2 — the non-monotonic pressure-radius characteristic of an inflating spherical elastic membrane, and the near-saturation of exhaled air — is standard textbook material in continuum mechanics and respiratory physiology respectively. Because no source was retrieved and verified during preparation of this document, no citation is given rather than an unverified one.

CITATION NEEDED — before any external circulation, the following should be supported by human-verified primary sources: (i) the hyperelastic membrane inflation characteristic, for which the Mooney-Rivlin and Ogden treatments of spherical membrane inflation are the standard references; (ii) the temperature and saturation state of exhaled air; and (iii) the gap statement in Section 1.3 regarding prior instrumented measurements of balloon burst pressure. No citation should be inserted without a human reading the source.

Appendix A. Data dictionary and file inventory

The complete data dictionary appears in Section 2.2 and the data-flow inventory in Section 2.3. Per-file validation results and the load-status log, including the two duplicate exclusions, are exported by the notebook as validation_report.csv and load_status_log.csv.

Files excluded as exact duplicates: sensor_qSPyn20BG36QUnuU0Mms.json (identical to sensor_CxSWwrpTqAwr5YoG19ev.json) and sensor_wpA9N1MvGEI0PueJWC3I.json (identical to sensor_d75GXJIxcHYbEOKzitHJ.json). Both were verified by MD5 content hash. Neither file was deleted or modified; both remain in the data directory untouched.

Appendix B. Per-run results

Selected run-level features for all 28 analysed runs. The full table, including start and rupture timestamps, ambient baselines and all derived features, is exported as run_level_features.csv.

Run ID	Inflation (s)	Burst gauge P (kPa)	Initial peak (kPa)	ΔRH (%)	ΔT (°C)	Samples
8ZjfiyTt5a…	39.4	5.10	5.30	32	0.7	1950
B4Ywn14vUy…	28.3	4.60	3.80	21	0.7	889
BC3s6CgDmk…	49.6	5.30	3.50	24	1.5	1657
CxSWwrpTqA…	28.3	5.00	4.00	12	0.4	1233
Gauqqh486Z…	28.3	5.90	3.60	8	0.2	860
IqtQTCa15V…	46.2	5.30	3.60	37	0.5	1473
J3OWMCoqua…	60.7	6.50	4.20	17	1.2	1876
KgfoeYAopk…	11.1	3.50	3.90	16	0.2	1044
MtUNDo3Kcd…	55.6	5.40	3.50	36	1.1	1926
OPIFS8NP0p…	28.3	5.50	4.90	22	0.3	1232
QhIJBf9OZg…	29.3	4.40	4.70	20	0.2	980
RwftwcifJ5…	45.5	6.30	4.20	24	0.7	1426
Tjb6MQ7b2k…	31.3	4.70	4.10	15	0.3	1266
VitpPf5VYh…	52.6	6.00	3.60	16	0.9	2613
Y5s8M68TBX…	61.7	5.20	3.20	30	1.0	2194
ZP55DFBUpj…	74.8	6.00	4.90	22	2.2	2110
d75GXJIxcH…	35.4	3.40	3.90	20	0.6	1202
eNslUJjuzo…	26.2	5.50	5.20	32	0.3	1057
eXg0HgvQjB…	39.4	4.50	4.50	13	1.4	1257
kKdULpYC2B…	53.6	4.50	3.00	36	2.2	1612

Run ID	Inflation (s)	Burst gauge P (kPa)	Initial peak (kPa)	ΔRH (%)	ΔT (°C)	Samples
lGGoUdFw8K…	62.7	5.10	3.80	27	1.0	1899
mfz4JTKmTy…	57.6	5.90	5.70	23	0.8	1920
mxL25LsgM7…	37.4	4.30	3.30	44	1.0	1288
pldTnbvGaa…	13.1	4.70	4.20	32	3.5	1956
qngL3jYz8k…	48.5	5.95	5.45	39	1.0	1383
ur11XTZQvh…	32.3	4.40	5.50	20	0.9	1198
vCQddO67yH…	26.3	2.90	5.10	21	0.1	1430
xctyYkbb77…	43.5	4.90	4.10	31	0.3	1615

Appendix C. Analysis environment and reproducibility

Component	Version / value
Python	3.11.14
NumPy	2.3.4
pandas	2.3.3
SciPy	1.15.3
Matplotlib	3.10.7
Random seed	42 (fixed; bootstrap intervals exactly reproducible)
Bootstrap resamples	10,000, percentile method, resampling runs
Notebook	balloon_analysis.ipynb — runs top-to-bottom from a clean kernel

The notebook fails loudly rather than silently on a wrong file count, an unexpected schema or header, a non-numeric value, a run in which no inflation is detectable, or a detected rupture that does not return to ambient pressure. Every number in this report corresponds to a printed output cell in that notebook. Raw data files were opened read-only and none was modified.

Appendix D. Artificial-intelligence use log

AI assistance was used throughout the analysis and drafting of this report. The table below records each material use, what was taken from it, and how a human verified the output. No data, results, citations or experimental context were generated by AI.

Purpose	What AI produced	Human verification
Challenge the research question	Identified that lung capacity cannot be derived from pressure-only data; proposed reframing to burst-pressure reproducibility	Confirmed by inspecting all 30 files for any volume or flow channel — none exists

Purpose	What AI produced	Human verification
Data-quality plan	Proposed content hashing, sampling-interval checks, plausibility bounds and a quantisation check	All checks executed in the notebook; duplicate finding independently reconfirmed with md5sum at the command line
Event-detection rule	Ambient baseline, threshold crossing, largest-drop rupture rule	Verified by the return-to-ambient assertion passing 28/28 and by the threshold sweep in Section 3.5
Analysis code	All loading, validation, feature-extraction, statistics and plotting code	Notebook restarted and executed end-to-end with zero errors; all assertions pass; every figure visually inspected
Statistical critique	Flagged the rate-duration denominator coupling and the row-level pseudo-replication risk	Coupling confirmed by the partial-correlation check in Section 3.4; unit-of-analysis choice documented in Section 2.2
Report drafting	Drafted all sections of this document from executed notebook outputs	Every reported number traced back to a printed cell output; no figure or statistic appears here that is not in the notebook

Explicit non-uses. AI was not asked to invent missing experimental context, to improve statistical significance, to remove inconvenient data, or to generate citations. Where context was unavailable — sensor model, calibration history, balloon specification, sensor placement — this report states the gap rather than filling it. Where a source could not be verified, Section 6 records CITATION NEEDED rather than supplying an unverified reference.