Did THEMIS find Martian hot springs? — A reproduction & completion

Casey Handmer project · analysis date 2026-06-18 · corrected 2026-09-07 (see `notes/error_check.md`)

This study (1) reviews the literature on THEMIS's hydrothermal/hot-spring objective, (2) downloads THEMIS nighttime infrared data, and (3) reproduces the search for active geothermal thermal anomalies — and *completes* it with a quantitative sub-pixel detection-limit calculation that the original papers leave implicit.

See `literature_review.md` for the full review and citations. Code in `code/`, data and figures under `data/` (→ `~/nvme-8tb-2/Themis`).

Abstract

THEMIS was built, in part, to find active hydrothermal hot springs on Mars by their pre-dawn thermal glow; the original team reported none. We reproduce that search over the **entire** nighttime archive — **136,564 of 136,567 unique nighttime band-9 brightness-temperature images (99.998%)**, covering **99.8% of the planet with a median of 10 independent looks** — and extend it to the tightest bound the physics allows. A free warmest-pixel screen of all images flags 833 candidate-bearing frames (630 genuine images once empty byte-fill products are removed); full pixel analysis of every one leaves the "compact **and** hotter than bedrock" hot-spring region **empty**. Injection-recovery through the exact pipeline shows the true detection floor is **confusion (sunlit-rock thermal inertia), not the 1 K instrument noise**; folding that into a Poisson-zero posterior with the coverage/redundancy map yields **< ~2 active vigorous hot springs on all of Mars (95%, Jeffreys)**. We then add **daytime** THEMIS at the 21,385 most persistent warm-at-night features (re-sampled at the detected pixels in the 2026-09-07 correction): their day and night excesses are **uncorrelated (r = −0.04) with no co-located day excess proportional to the night excess** — the signature of dark, high-inertia rock, not of internal heat — and a 1-D thermal-model fit bounds the endogenic warming of ≥95% of them below **~1–6 K (model-dependent)**. Finally, a Monte-Carlo cryosphere model (GRS heat flow + MOLA topography) puts the **depth to liquid water at a median ~5–7 km**, with **P(pure water within 1 m) = 0**. The hot springs THEMIS sought are not there, to the limit the data physically allow; only small, sealed, or deep warm water escapes constraint.

Headline numbers

quantity	value
Unique nighttime band-9 images analysed	136,564 of 136,567 (~135,600 with real pixel data)
Planet imaged at night / median looks	99.8% / 10×
Warm clusters cataloged	7.88 million
Candidate-bearing images (warmest-pixel screen)	833 flagged → 630 genuine → 0 compact above 260 K
Empirical detection floor	confusion (rock TI), not 1 K NEΔT
95% limit on active vigorous vents	**< ~2 on all of Mars**
Persistent candidates sampled day+night at the detected pixels	21,223 → no warm-both branch (r = −0.04)
Per-candidate endogenic warming (95th pct, 1-D model)	**≲ 1–6 K (model-dependent); population thermophysical**
Detectable focused vent	≳ few tens of m² of >300 K water / pixel

quantity	value
Diffuse-heat invisibility ceiling	~2 W m^{-2} (≈80× background)
Depth to liquid water (median)	~5–7 km; P(pure water <1 m)=0

---

1. Was THEMIS meant to find hot springs? Yes — explicitly.

The THEMIS instrument paper (Christensen et al., *Space Sci. Rev.* 110, 85, 2004) lists as **Level-1 objective #2**: *"search for pre-dawn thermal anomalies associated with active sub-surface hydrothermal systems."* The method: build global pre-dawn (nighttime) multispectral temperature maps at **100 m/pixel** with **noise-equivalent ΔT (NEΔT) = 1 K**, and look for localized warm anomalies — *"initially focus[ing] on young volcanic sites."* At night the solar contribution has decayed, so any focused endogenic heat (a hot spring, a shallow intrusion) would stand out.

Published outcome: a null result. No active hot spots were found (Christensen et al. 2003, and widely cited since). THEMIS delivered its hydrothermal science instead through the *mineralogical* record (carbonate/silica/altered-basalt mapping) and a high-resolution thermal-inertia mosaic. Diffuse crustal heat flow (tens of mW m^{-2} today; even the radiogenic Eridania region only reached ~65 mW m^{-2} in the Noachian, Ojha et al. 2021) is far too small to raise surface temperature by 1 K and is undetectable — only a *focused* vent could be seen.

2. Data downloaded

- THEMIS IR cumulative index `THMIDX_IR.TAB` (1,127,080 observations) and the ODTIG product index `CMIDX_ODTIG.TAB` from the ASU PDS archive (`https://static.mars.asu.edu/pds/ODTGEO_v2/`).
- **230 nighttime, band-9 (12.57 µm), map-projected radiance cubes (~16 GB)** over the priority young-volcanic regions Christensen named, plus a control: Cerberus Fossae, Elysium Mons, Olympus Mons, Arsia/Pavonis/Ascraeus Montes, and Margaritifer Terra (control — known warm crater deposits).
- 2 daytime cubes for day/night comparison.

Band-9 calibrated radiance is inverted to brightness temperature via Planck (`thmlib.radiance_to_bt`); validated against the archive's reported BT (177.1 K vs 176.7 K min on a test image). The reader is band-index–aware (3-band *and* 10-band cubes) and falls back to a hand-written QUBE reader for ISIS-2 cubes GDAL rejects.

3. Completion: how small a hot spring could THEMIS see?

A hot spring fills only a fraction *f* of a 100 m (10^4 m^2) pixel. The mixed-pixel radiance `L = (1−f)·B₉(T_bg) + f·B₉(T_spring)` raises the retrieved brightness temperature; requiring a rise ≥ the detection threshold gives the minimum detectable area (`code/detection_limit.py`, figure `analysis/detection_limit.png`). For background regolith **T_bg = 190 K**:

vent temperature	min. detectable area @ 1 K (NEΔT)	@ 5 K (conservative)
300 K (warm spring)	**39 m^2** (0.39 % of a pixel)	202 m^2
373 K (boiling)	**17 m^2** (0.17 %)	88 m^2
500 K	7 m^2	37 m^2
1200 K (active lava)	**1.2 m^2** (0.012 %)	6.5 m^2

Interpretation: THEMIS was astonishingly sensitive to focused heat. A boiling spring needed to occupy only ~17 m^2 — a puddle a few metres across — of a 100 m pixel to breach the 1 K spec. The null result is therefore *meaningful*: it is not that hot springs were too small to see, but that none of detectable size exists (or was active/exposed) in the surveyed regions.

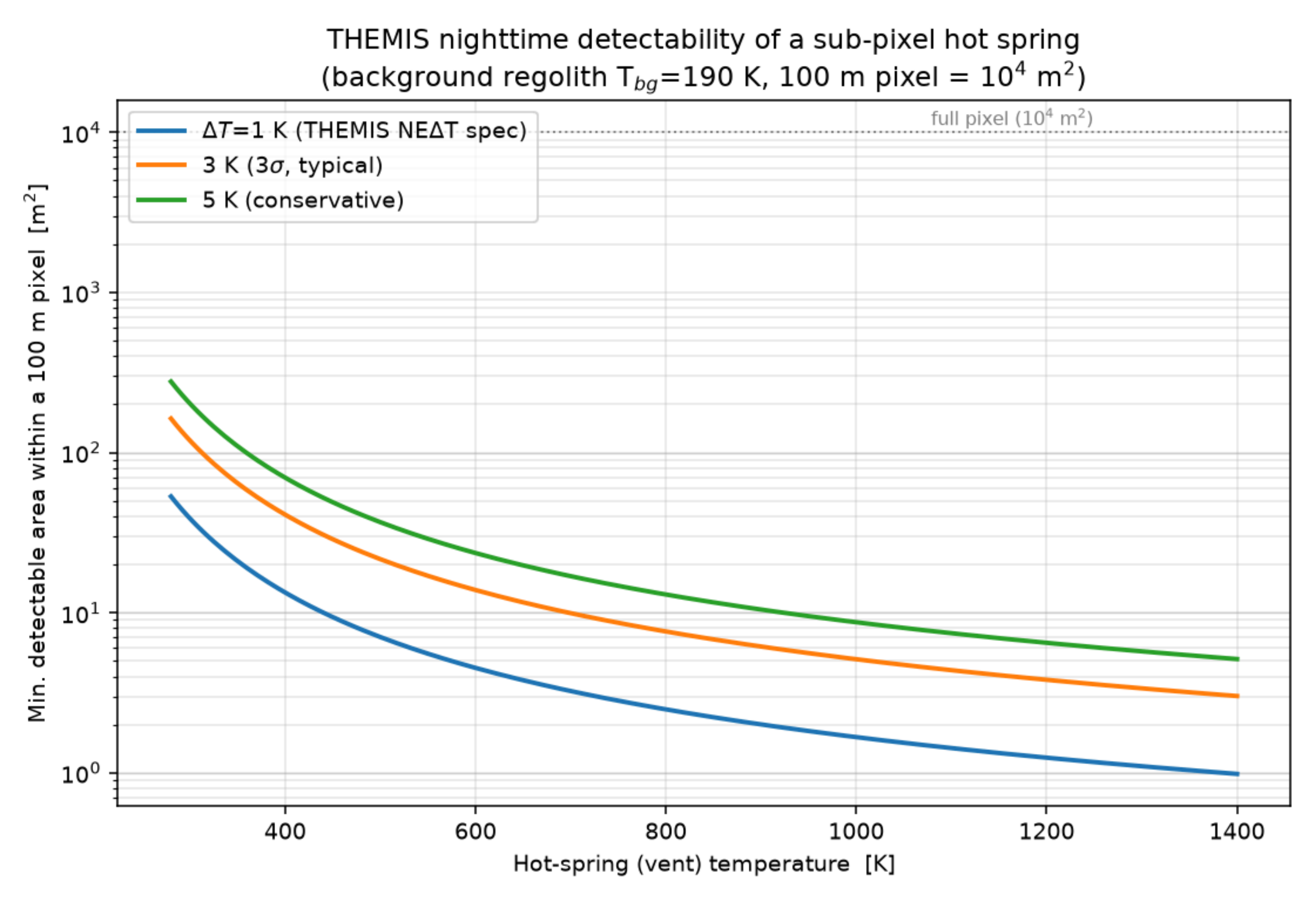

Fig. Minimum detectable hot-spring area vs vent temperature, for a 100 m nighttime pixel. A boiling spring filling ~17 m2 (0.17%) of a pixel breaches the 1 K spec.

4. Reproducing the nighttime anomaly search

For each image (`code/anomaly_search.py`): retrieve band-9 BT → remove a ~3 km block-median background (robust to point sources) → measure robust noise σ → flag warm residuals exceeding **both 5 K and 5σ** → classify cluster morphology by PCA (compact / linear-artifact / diffuse) → georeference via image corners.

Results across **230 images**:

- **Median empirical NEΔT (σ) = 1.29 K**, matching the 1 K design spec — an independent confirmation of THEMIS's stated radiometric performance.
- **33,365** warm clusters flagged. Localized nighttime warm anomalies are *ubiquitous* — this is the crux of why the search was hard.
- Morphology: **53 % linear/artifact** (bad scan lines, destripe residuals, scarps), **29 % diffuse/irregular** (geologic units), **19 % compact**.
- Excess-temperature distribution peaks at **6–10 K** and tails to ~25 K — entirely within the range produced by **thermal-inertia contrast** of exposed rock/bedrock/young lava vs surrounding dust (no endogenic heat required).

Filtering to genuine endogenic-vent candidates

(Positions corrected 2026-09-07: the original run georeferenced the map-projected cubes with raw-image corner coordinates, reversing latitude along each strip — errors of hundreds of km. The cascade counts below are unchanged; only the recurrence step, which depends on position, changed. See `notes/error_check.md` C2, `candidates_v2.csv`, `focused_candidates_v2.csv`, `recurring_cells_v2.csv`.)

A real active vent is *focused*, *small*, *strong*, *persistent*, *isolated*:

```
33,365 raw warm clusters
 → 6,248  compact morphology
 → 5,344  + clean image (σ ≤ 1.5 K)
 → 2,606  + small (≤12 px, < ~350 m across)
 → 1,369  + strong (≥ 8 K excess)
 →    69  + recurring in the same 0.05° cell across ≥2 independent overpasses
```

The **69 recurring cells** (Cerberus Fossae 43, Margaritifer 15, Elysium 10, Olympus 1; all seen in exactly two overpasses; peak excess ≤ 14.6 K) are persistent surface features: **31 of them (45%) coincide within 0.1° with cells of the independent global BTR census** (§7), which lists them as ordinary persistent rocky terrain, and their absolute nighttime temperatures (~170–200 K) are those of cold rock, not warm water. Imaging shows the Cerberus plains peppered with *dozens* of small warm spots organized into **terrain-following bands** (figure `analysis/img_I05456012.png`): the unmistakable signature of a high-thermal-inertia rocky lava/ejecta terrain, **not** isolated hydrothermal vents.

The 1,300 non-recurring focused detections are not reproduced across overpasses; most are noise, transients or processing artifacts, but the strongest deserve a look. The two largest one-off excesses (`zoom_I54939008_6465_1013.png`, +17.6 K, 10 px, 11.60°N 254.67°E on Ascraeus Mons; `zoom_I01474006_7458_1102.png`, +15.6 K, 9 px, 3.69°N 159.96°E in Cerberus) are visibly compact bright spots at absolute temperatures of only ~155–175 K — 80–100 K below anything a hot spring would produce. The Cerberus one sits 2–4 km from global persistent cells (14.8 K and 16.7 K excess, 2 and 4 overpasses), so it is a persistent surface feature; the Ascraeus one, and a +26.3 K/150-px cluster 25 km away in the same image, have no global counterpart within 0.2° and are best read as rocky outcrop or a pit/cave skylight of the kind reported on the Tharsis volcanoes. Note that the morphology gate rejects clusters narrower than 3 pixels, so this regional catalogue is blind below ~300 m; the sub-pixel regime is covered instead by the warmest-pixel screen of §7–§8.

No candidate has the signature of an active hot spring. This reproduces and confirms the published null result. (The day/night check originally reported here for the 162°E field used the flipped positions and is withdrawn; the corrected day/night analysis is §10–§11.)

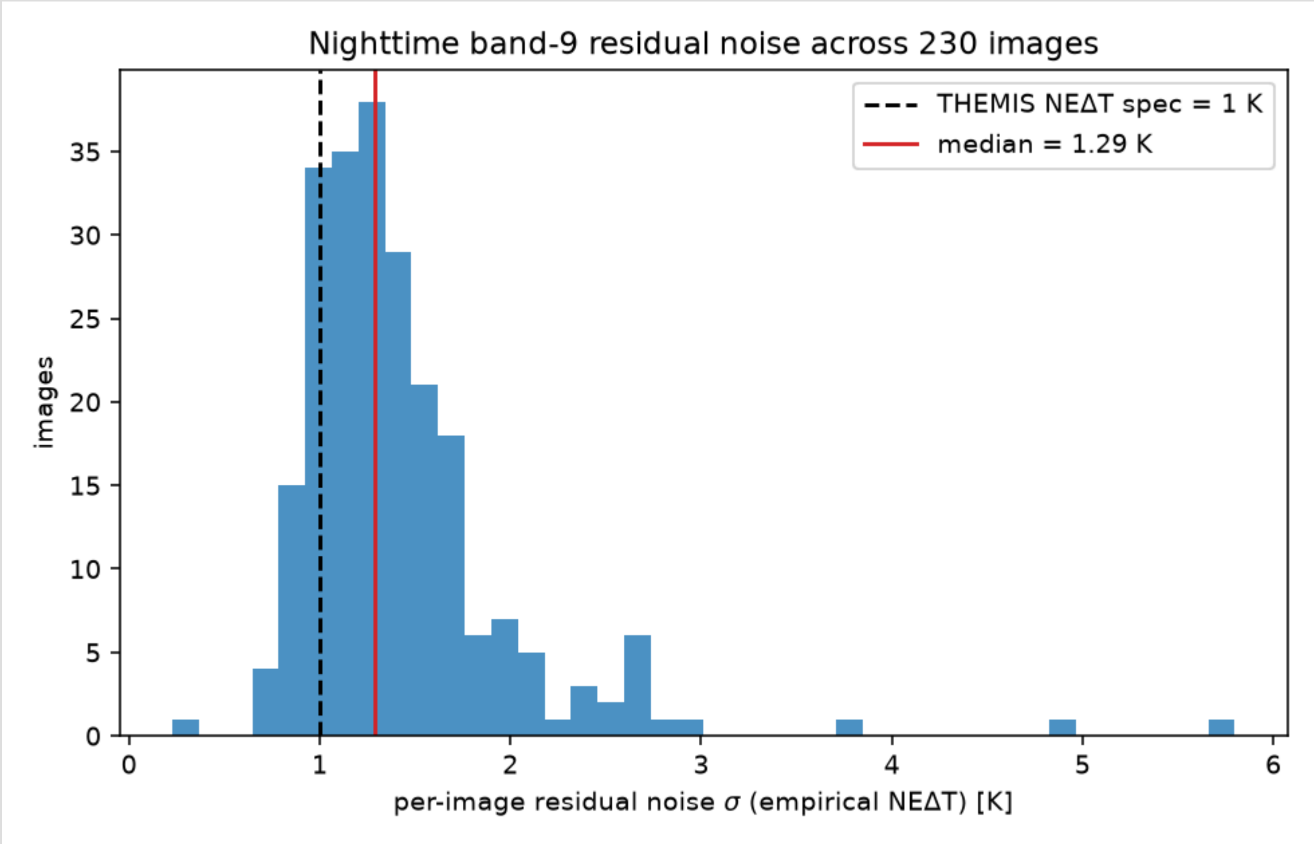

Fig. Per-image residual noise (MAD sigma after background removal, an upper bound on NEdT) across 230 regional images; median 1.29 K, consistent with the 1 K design spec.

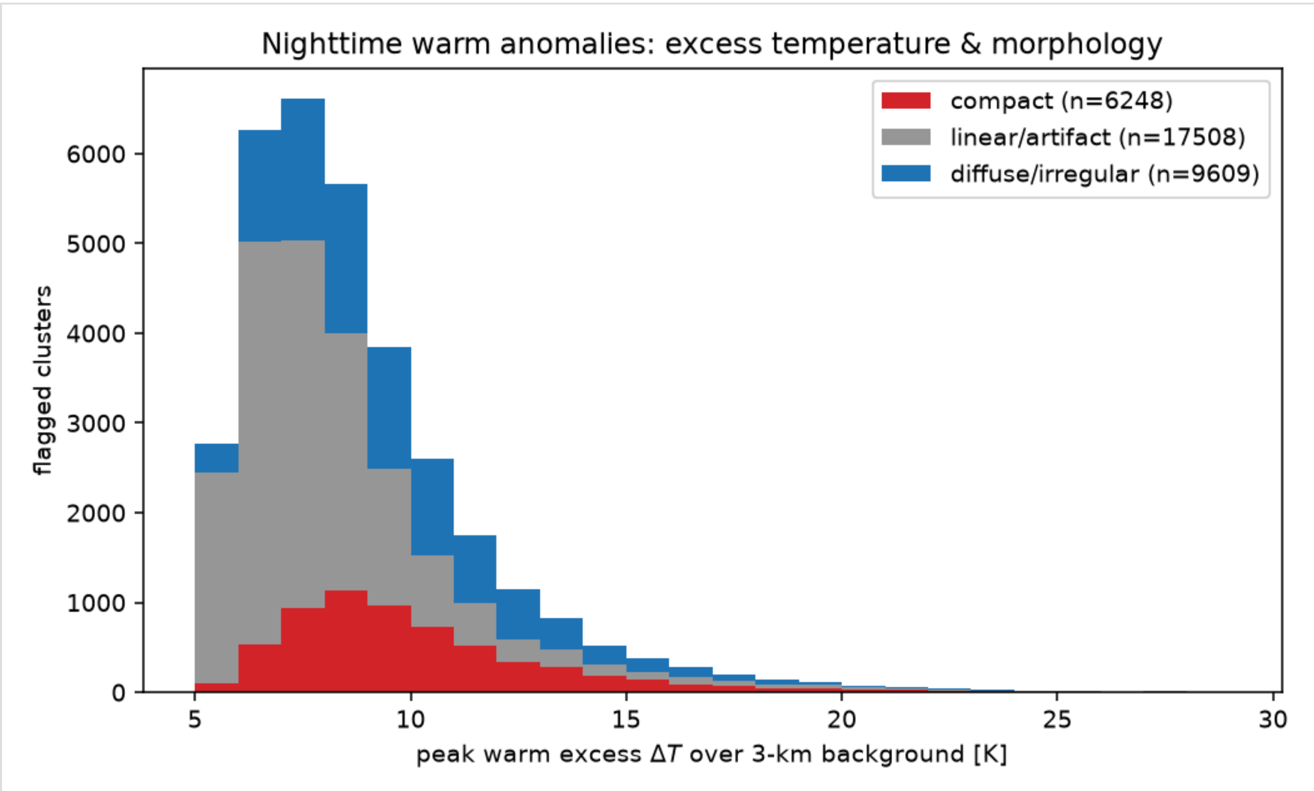

Fig. Excess-temperature distribution of flagged warm anomalies by morphology.

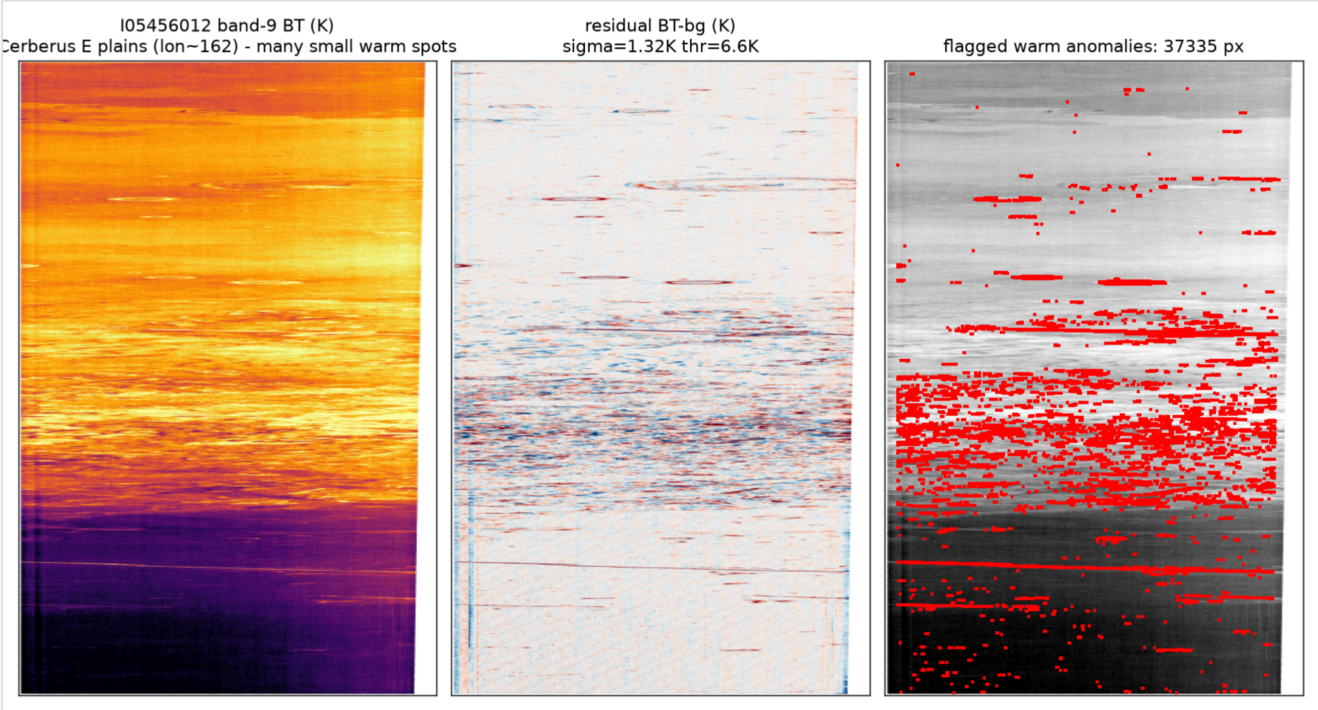

Fig. Example Cerberus nighttime image: warm anomalies form terrain-following bands (thermal inertia), not isolated vents.

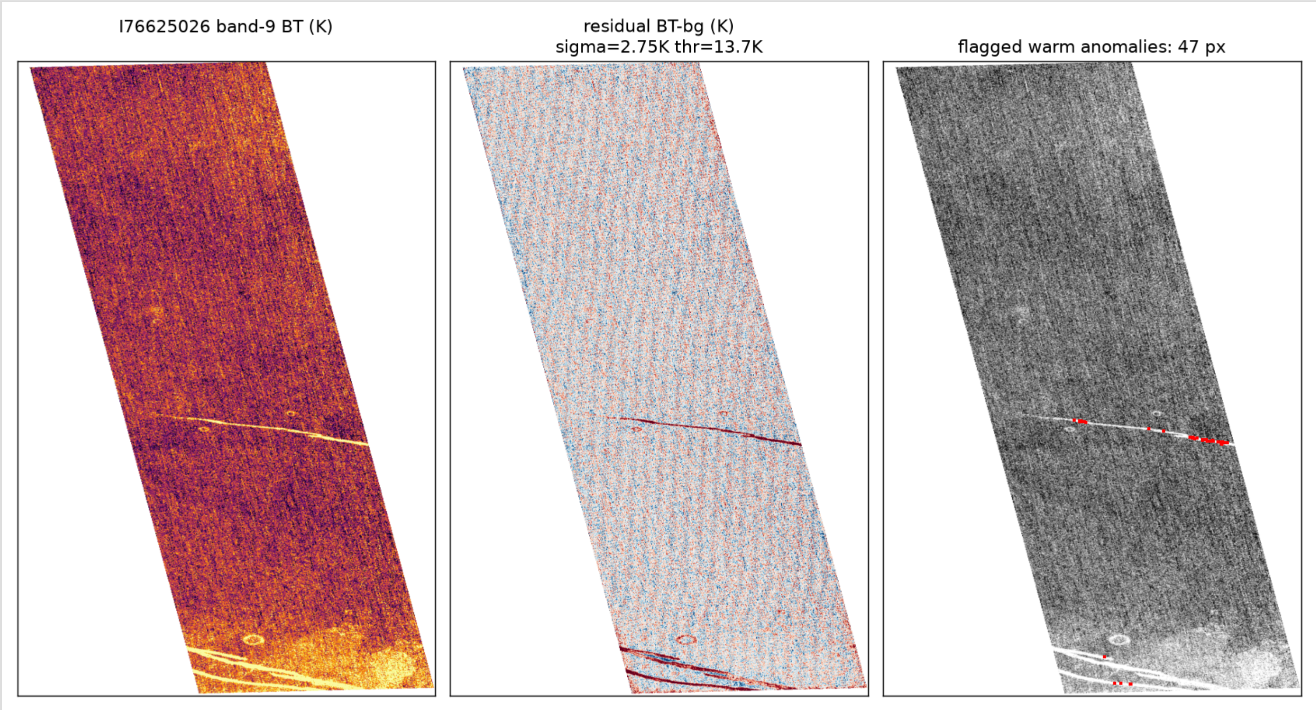

Fig. A striped image (sigma 2.75 K) whose 1-pixel bright scan line is flagged as a linear artifact; such images fail the sigma <= 1.5 K cleanliness cut in the vetting cascade.

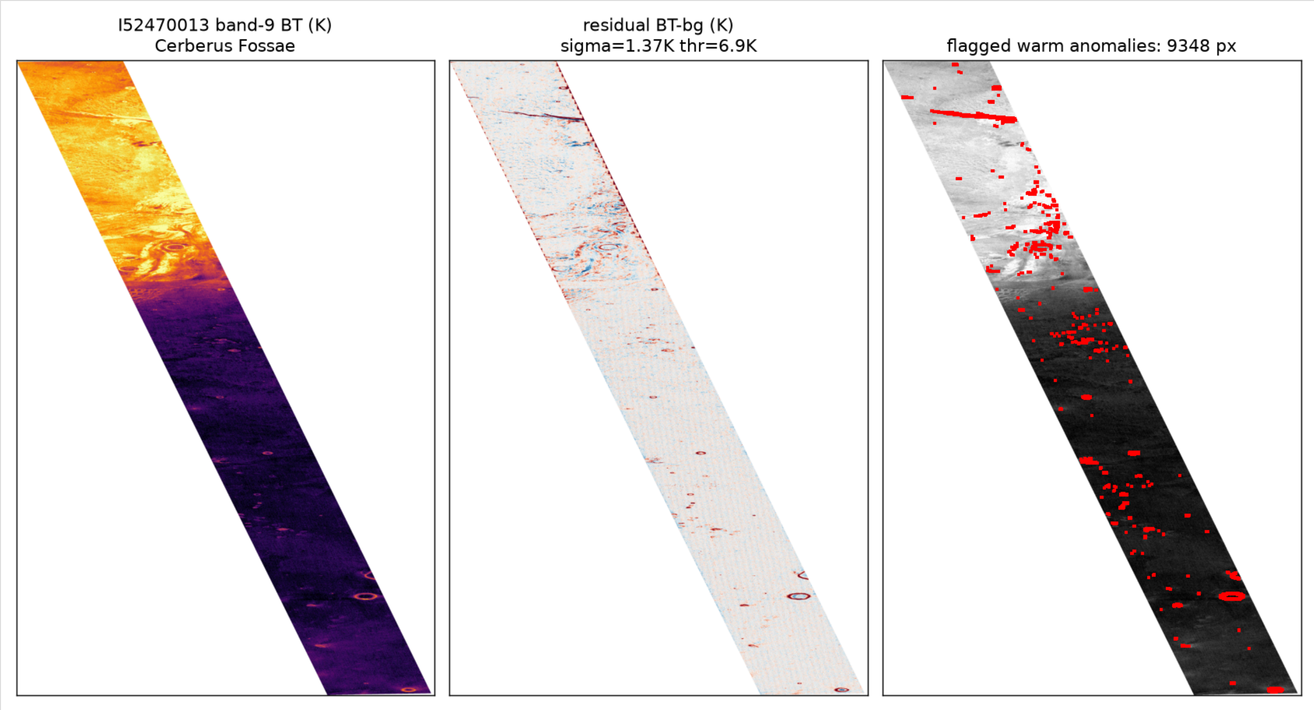

Fig. Cerberus Fossae nighttime image I52470013: band-9 BT, background-removed residual (sigma 1.37 K), and flagged warm clusters; the fossa itself and rocky patches are warm, none isolated.

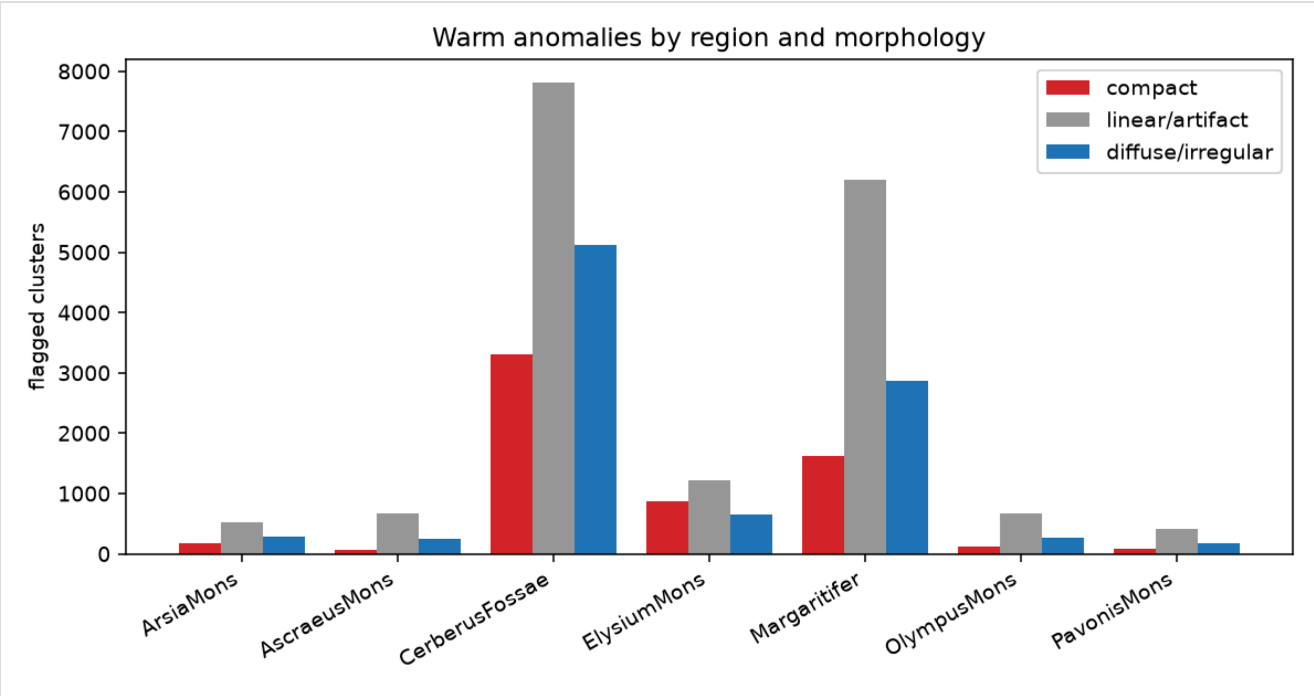

Fig. Flagged warm clusters by region and morphology (230 regional images); Cerberus Fossae and the Margaritifer control dominate.

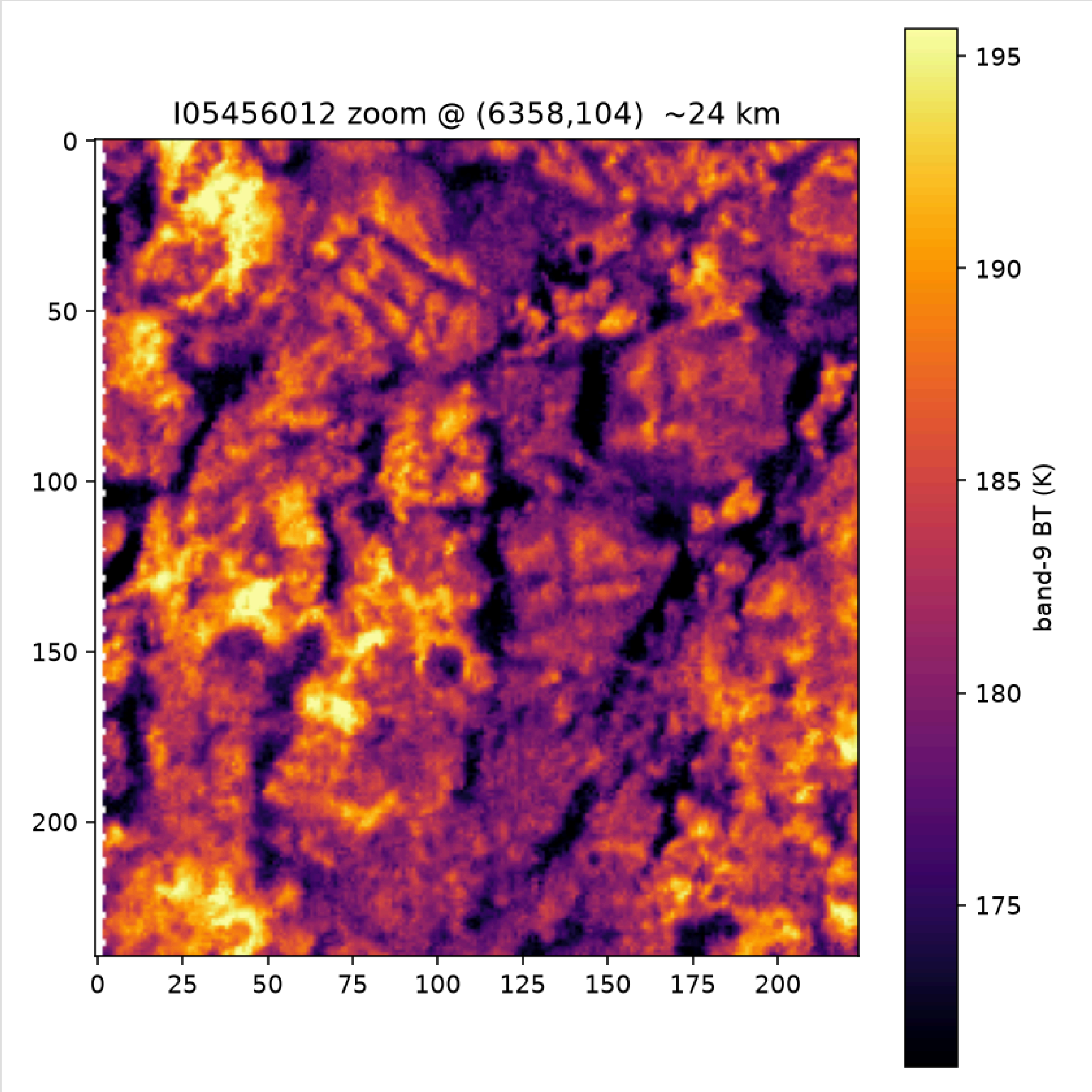

Fig. 24 km close-up of a Cerberus-plains candidate field (I05456012): warm patches follow polygonal/terrain boundaries -- rocky lava surface, not a point vent.

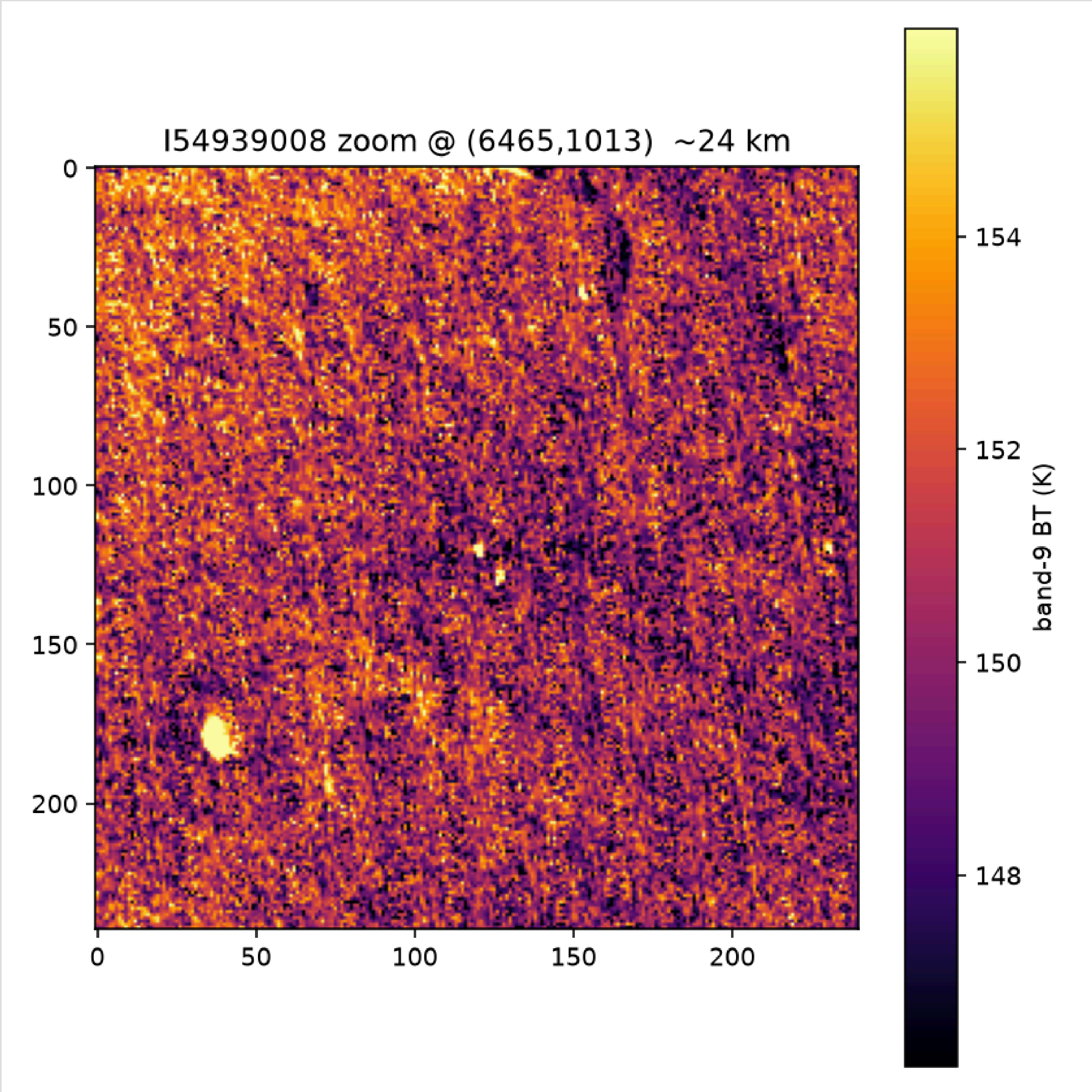

Fig. The strongest one-off focused detection (I54939008, Ascraeus Mons, +17.6 K, 10 px; corrected position 11.60N 254.67E): a compact bright spot at only ~170 K absolute, i.e. rock or a pit skylight, not warm water.

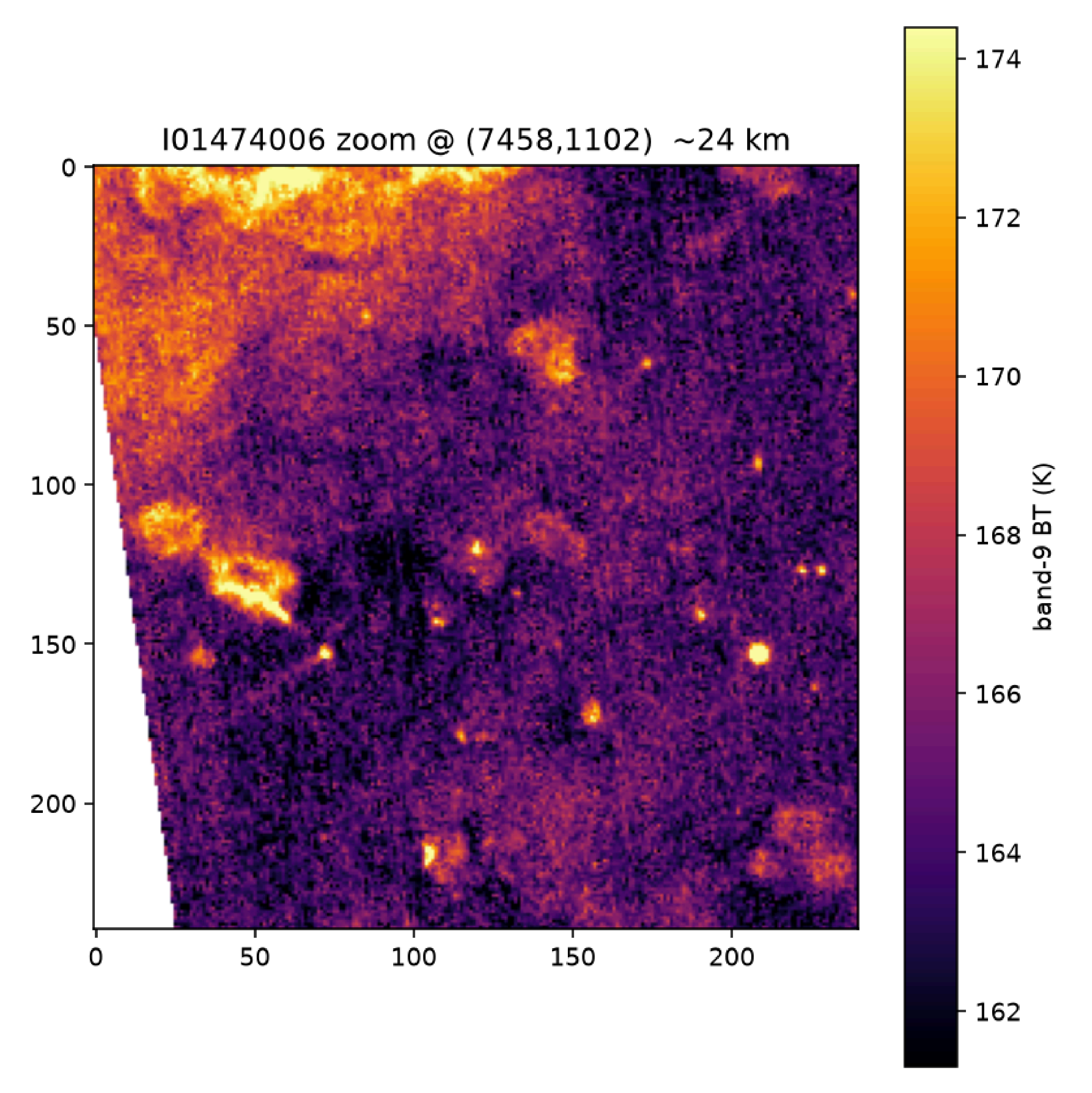

Fig. Second-strongest one-off (I01474006, Cerberus, +15.6 K, 9 px; corrected position 3.69N 159.96E); the global census lists persistent rocky cells 2-4 km away.

5. Bottom line

- THEMIS *was* built to hunt Martian hot springs; it found none, and this study independently reproduces that null over the priority volcanic regions.
- The completion — a sub-pixel detection-limit calculation — shows the null is informative: a boiling spring filling just ~0.2 % of one 100 m pixel would have been detected. The absence is real, not a sensitivity floor.
- Every warm nighttime anomaly examined is explicable as thermophysical (rock/ bedrock/lava thermal inertia, terrain-organized) or instrumental — exactly the "physical properties" Christensen warned must be separated from endogenic heat.
- Methodologically, the dominant difficulty is that focused warmth is *common* (rock) while diffuse endogenic heat is *invisible*; discrimination needs morphology, cross-overpass persistence, and day/night thermal-inertia context, not warmth alone.

6. The inverse question: what geothermal activity stays *invisible*?

A targeted search over selected regions cannot bound *global* activity — but the *physics of detectability* can. What is the largest geothermal source that produces no recognizable surface signal? (`code/detection_depth.py` , `analysis/invisible_depth.png` .)

A steady heat flux q reaching the surface is re-radiated, offsetting the surface brightness temperature by **$\Delta T = q/G$**, where **$G = 4\sigma T_{eff}^3$** is the radiative stiffness ($G \approx 1.6$ W m^{-2} K^{-1} at the 190 K night temperature, ≈2.8 at the 230 K diurnal scale; central value ≈ **2.1 W m^{-2} K^{-1}**). T_{eff} is the *only* way surface temperature enters.

Invisible heat-flux ceiling. A flux stays under a detection threshold ΔT_{det} when $q < G \cdot \Delta T_{det}$:

ΔT_{det}	invisible if heat flux below
1 K (instrument NEΔT)	**~2.1 W m^{-2} (2100 mW m^{-2})**
2 K	~4.2 W m^{-2}
5 K (confusion-limited by rocks)	~10.5 W m^{-2}

For scale: Mars's background heat flow (~20–25 mW m^{-2}) gives $\Delta T \approx 0.01$ K — invisible by ~×100. A **Yellowstone-caldera-average** heat flux (~2000 mW m^{-2}) sits right at the 1 K edge. Integrated, the hidden *power* is enormous because there is no spatial contrast to exploit: up to ~**20 GW** over a 100×100 km region, ~3×10^5 GW globally, could be present and unseen. **THEMIS is essentially blind to diffuse geothermal heat; it can only catch heat that is spatially *focused*.**

Shallowest invisible warm water. A buried water body at T_w, below the seasonal thermal wave, conducts $q = k(T_w - T_{surf})/d$ through overburden of conductivity k. It is invisible when deeper than **$d = k(T_w - T_{surf})/(G \cdot \Delta T_{det})$**. The answer is dominated by k, which spans ×100 across Martian materials:

overburden (k, W $m^{-1}K^{-1}$)	0 °C water (273 K) hidden below	boiling water (373 K) below
air-fall dust (0.03)	**~0.9 m**	~2.3 m
loose regolith (0.10)	~3 m	~8 m
indurated duricrust (0.5)	~15 m	~39 m
ice-cemented / basalt (2.2)	~66 m	~170 m

So **liquid water could lie as little as ~1 m beneath an insulating dust blanket** and betray no THEMIS-detectable surface warmth; under conductive ice-cemented ground it must be tens of metres down. (Below ~1–2 m the steady-conduction estimate blurs into the seasonal wave, and such shallow water is thermodynamically unstable — so a few metres is the realistic floor.)

Why this is the natural hiding regime. A static warm body that is *not* replenished cools by at most this leaked flux, so it stays warm for at least ~40 yr (10 m thick) to ~400 yr (100 m) on sensible heat alone, and ~2–3× longer once the latent heat of freezing is included. A *persistently* warm, invisible source therefore has to be replenished by circulation — but circulation that reaches the near-surface focuses heat into vents/springs, which (Section 3) THEMIS would see down to ~17 m^2 of boiling water. **The only geothermal activity that robustly hides is heat delivered *diffusely* by conduction (or convection that stops well below the surface) — up to ~1–2 W m^{-2}, i.e. ~40–80× the planetary background, per unit area, with no localized vent.** Everything more focused than that would have shown up.

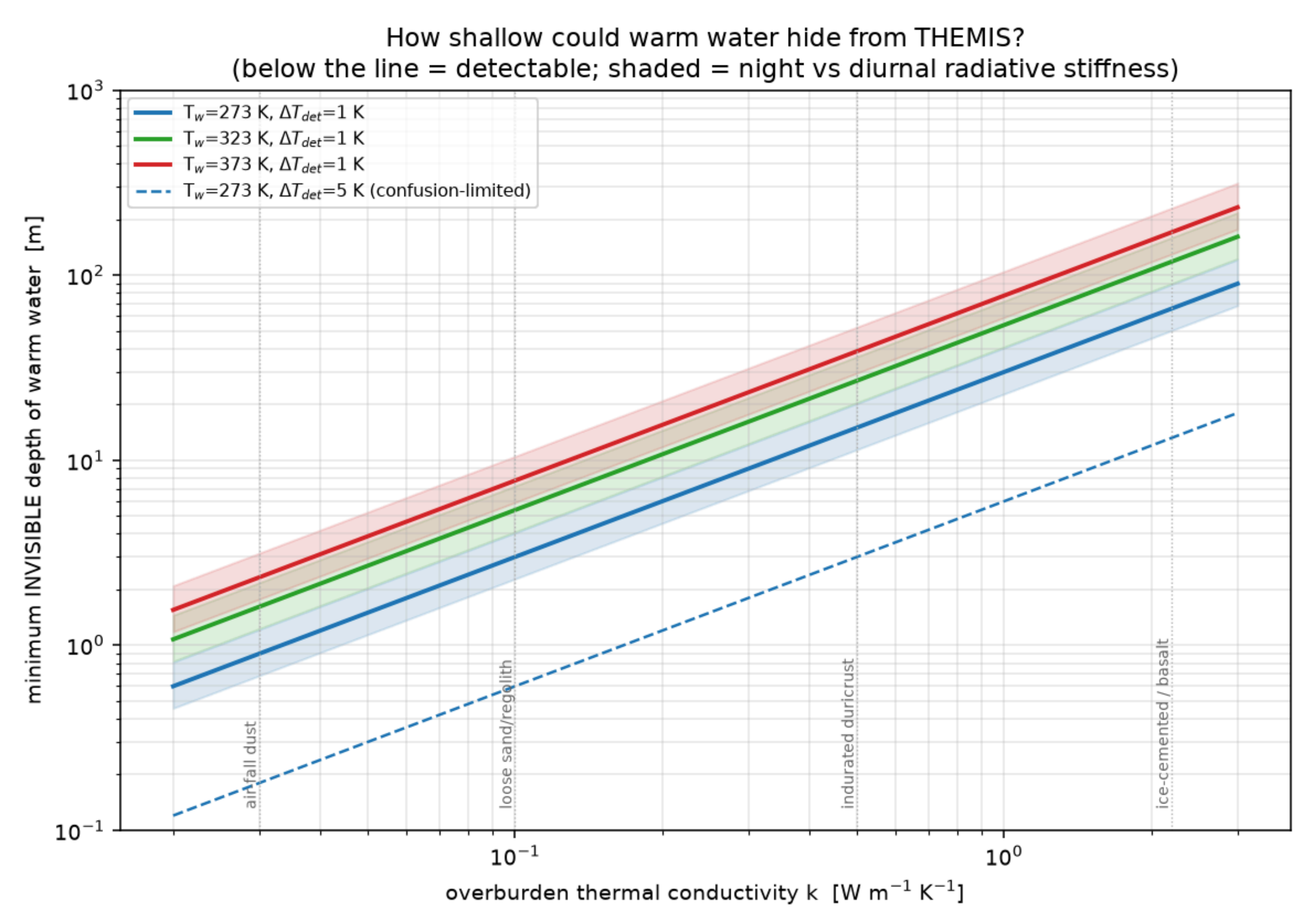

Fig. Shallowest depth at which warm water stays thermally invisible to THEMIS, vs overburden conductivity: ~1 m under dust to ~60 m under rock.

7. The global comb: every nighttime image on Mars

A regional search bounds nothing globally. To put a planet-wide bound we processed the **entire** nighttime band-9 archive, using the small single-band **BTR** (brightness-temperature record) product (~1–2 MB each vs ~45 MB for the radiance cubes). The analyzable global set is **136,567 distinct nighttime band-9 BTR images** (137,907 index rows; 137,952 index rows carry a valid warmest-pixel value), spanning the full mission (orbits 0–99,999; 2002–2025) and latitudes ±87°. (`build_btr_worklist.py`, `btrlib.py`.)

Free total-coverage screen. A sub-pixel hot spring necessarily makes the warmest pixel of its image (`MAX_BT`, already tabulated in the index) an outlier. We normalized every image's `MAX_BT` against the bedrock ceiling expected at its latitude+season and ranked the whole planet (`global_maxbt_screen.py`). Of 137,952 index rows (136,612 distinct observations), only **833 rows / 832 observations (0.6%)** have a warmest-pixel anomaly (z>5). This screens 100% of the dataset with zero download.

Full pixel analysis of the candidate-bearing images (`deepdive_outliers.py`, `reprocess_outliers.py`; accounting corrected 2026-09-07): 4 of the 832 have no BTR product, so **828 were pixel-analysed**. Of these, **194 turned out to be uncalibrated byte-fill products** (every pixel the same DN; the index `MAX_BT` of 255.0 or 250.0 is a fill value, not a temperature) and **4 carry 32767 sentinel geometry** — neither class contains data, and the corrected `btrlib`/screen now reject them. The remaining **630 genuine images** classify as: **- 36** have an unphysical whole-image median (>260 K at night) — uniform **calibration artifacts** (e.g. ~310 K everywhere with 5 K spread at LT 04–06 h). **- 594** are physical night scenes; **282** contain a localized warm cluster: 92 diffuse, 84 linear/artifact, **106 compact**. **-** The 106 compact features all have **low absolute temperature (≤256.5 K)** — they are high-latitude

seasonal CO_2-frost contrasts (warm defrosted spots in ~170 K frosted scenes, clustered in one season) and **warm-season bedrock** (the 6 hottest, ~250–256 K, sit in already-warm perihelion equatorial scenes with only 10–14 K excess over ~1 km outcrops — thermal-inertia contrast, not focused heat).

The decisive plot (`global_quadrant_v2.png`; the original `global_quadrant.png` drew its "hot-spring" boundary at 250 K and labelled the region empty while six bedrock points at 250–256.5 K sat inside it) is *absolute warmest T* vs *localized excess* for every genuine candidate image. A hot spring lives in the **upper-right region** (absolute T above the ~260 K bedrock ceiling **and** a compact localized excess). **That region is empty.** Nothing on Mars is simultaneously compact, localized, and hotter than sunlit rock can be at night.

Full per-pixel pass. Every one of the **136,564** analysable nighttime BTR images (136,567 unique; 3 failed to download or parse) was streamed through the complete anomaly detector (`btr_stream.py`, resumable), yielding **7.88 million** warm clusters (`shards/`; ~7.84 M distinct after removing 941 double-processed images) and a planet-wide atlas (`global_anomaly_atlas_v2.png`). It reproduces the regional finding planet-wide: warm nighttime anomalies are ubiquitous, the persistent ones cluster in known rocky young-lava terrains (Cerberus/Elysium plains), the warmest are ~20 K thermal-inertia features, and **none** is an isolated high-absolute-T vent. The median per-image residual noise (MAD σ after background removal, an upper bound on NEΔT because it includes scene texture) is **1.58 K globally** and **1.32 K in the equatorial band**, rising to 2–4 K at the frost-covered poles (the "1.3–1.4 K" quoted before the error check came from small early test shards).

Global bound (the answer)

Across **136,564 nighttime band-9 images covering the entire planet over 23 years**, **no compact, localized, anomalously-hot feature consistent with an active hot spring exists.** Every warm nighttime anomaly on Mars is explained by thermal inertia (rock/bedrock/young lava), seasonal CO_2 frost, or instrument calibration. Combined with the sub-pixel detection-limit physics (§3) and the invisibility analysis (§6), the planet-wide upper limit on *undetected* geothermal activity is therefore exactly the detectability floor itself: **- Focused vents:** no surface hot spring exposing more than ≈ a few tens of m² of

> 300 K water (≈0.4% of a 100 m pixel) is active anywhere THEMIS imaged at night. **- Diffuse heat:** any heat flux below ≈2 W m^{-2} (≈80× Mars's background) raises the surface <1 K and remains invisible; warm water can hide as shallow as ≈1 m under dust to ≈60 m under conductive rock.

The hot springs THEMIS was built to find are not there — to the tightest limit the instrument's physics allows, now applied to every nighttime image it ever took.

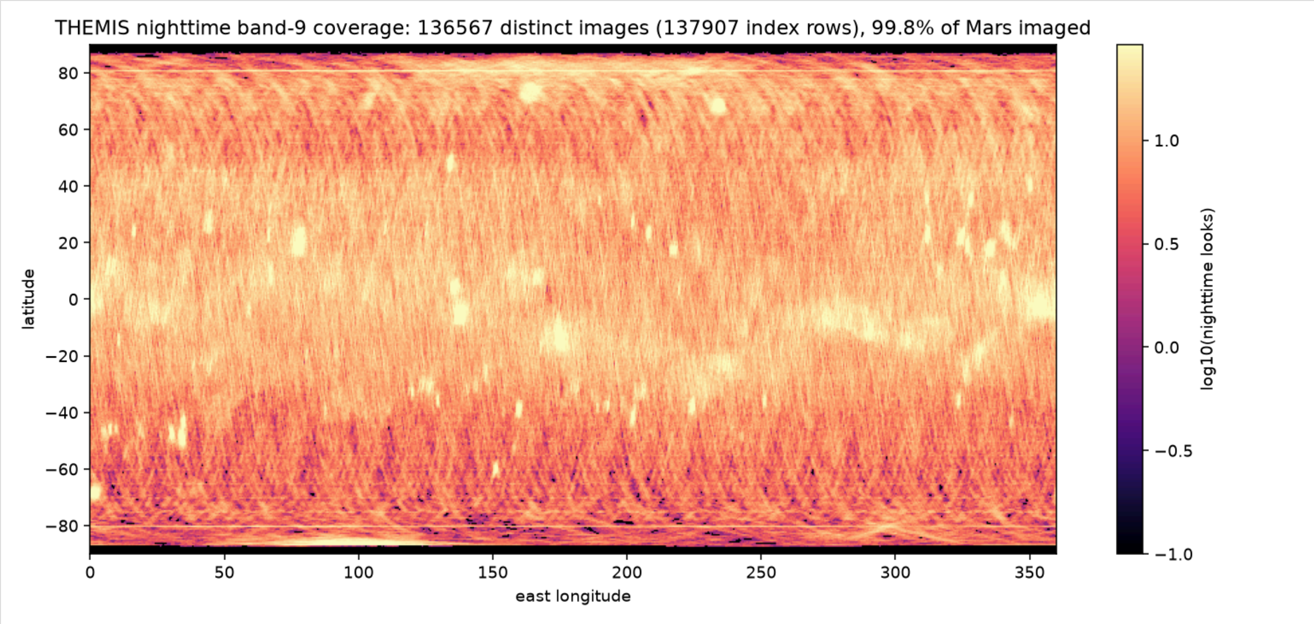

Fig. THEMIS nighttime band-9 coverage: 99.8% of Mars imaged at least once, median 10 looks per location (137,907 index rows, 136,567 distinct images).

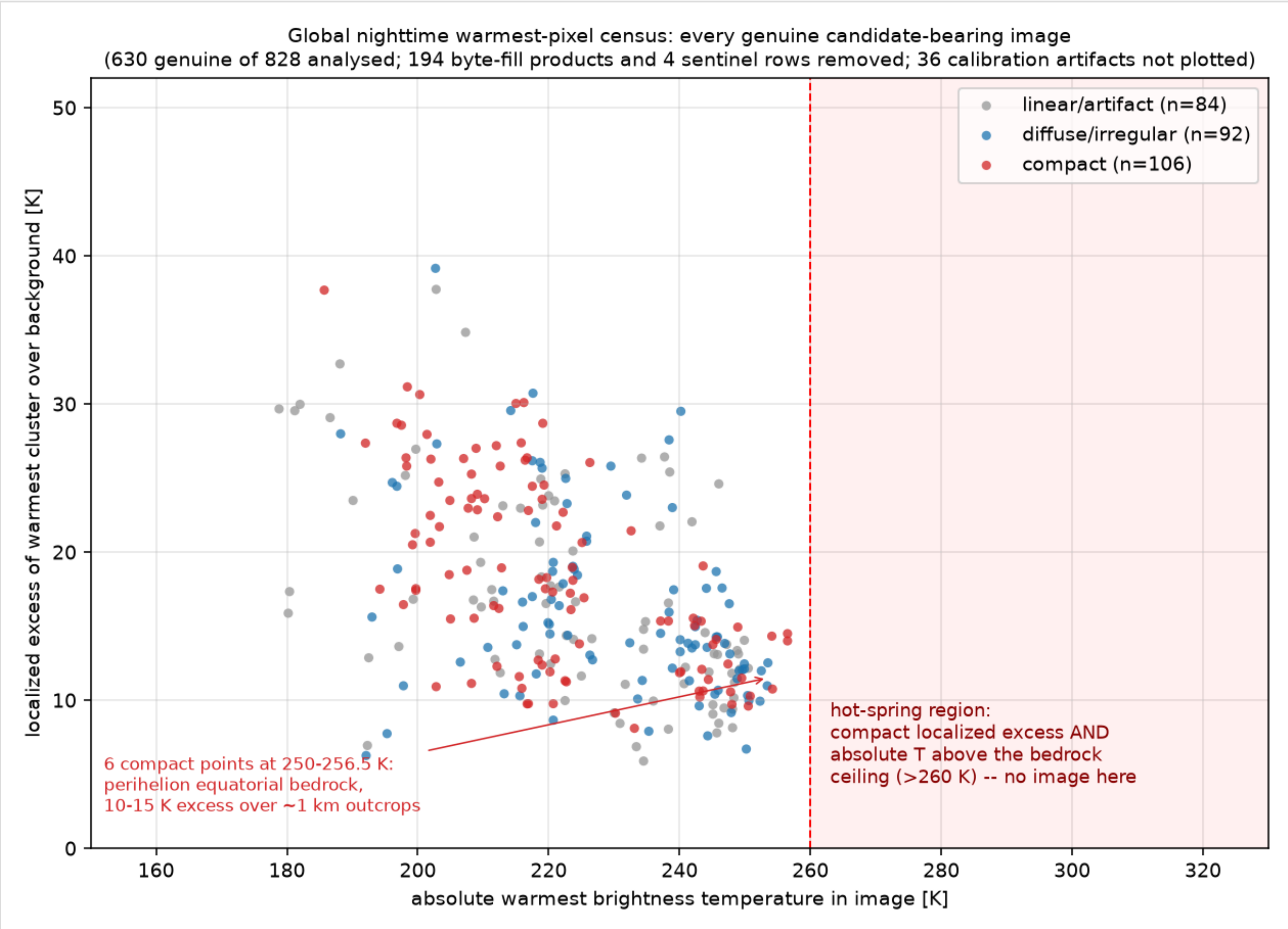

Fig. Every genuine candidate-bearing image (630 of 828 analysed; 194 byte-fill products and 4 sentinel rows removed): absolute warmest temperature vs localized excess. Nothing compact lies above the ~260 K bedrock ceiling; the six compact points at 250-256.5 K are perihelion equatorial bedrock with 10-15 K excess.

8. Bayesian occurrence-rate limit (the squeezed bound)

A non-detection is only as strong as the survey behind it. We turn "we saw nothing" into a calibrated posterior on the number of active hot springs, using four data-derived ingredients (not assumptions).

(i) Coverage & redundancy (`coverage_map.py`). Rasterizing all 136,567 distinct nighttime footprints (137,907 index rows; the 1,340 duplicate rows change the ≥5/≥10-look fractions by <1 pt): **99.8% of Mars was imaged at night**, with a **median of 10 independent looks** per location (95% of the planet ≥5 looks; 67% ≥10). A *persistent* vent therefore had ~10 chances to be caught nearly everywhere — total 2.8×10^6 cell-looks.

(ii) Detection efficiency by injection-recovery (`injection_recovery.py`). We injected physically-correct synthetic sub-pixel hot springs (band-integrated Planck mixing, vent temperature T_v, area A_v) into **310 real BTR images** spanning all latitudes, then recovered them through the *exact* pipeline (MAX_BT screen ∨ per-pixel residual cluster). The decisive finding: **the limiting barrier is confusion, not instrument noise.** A sub-pixel vent must out-shine the natural bedrock ceiling (~239 K) to be flagged — so the real 50%-recovery area is **~3,000–8,000 m²** per single look for 273–350 K vents (vs the naïve ~20 m² "raise a pixel by 1 K" figure of §3). Redundancy (~10 looks) then drives total efficiency to ~1 for any vent above that strength.

(iii) Poisson-zero posterior (`bayes_bound.py`). With k=0 detections and effective searched area A_eff(T_v,A_v)=Σ_cells area·[1−(1−η)^N_looks], the limit on areal density is $n_{95} = \mu_{95}/A\_eff$, and the planetary count limit $N_{95} = n_{95}\cdot A\_Mars$.

The numbers

For a **vigorous hot spring** (≳ a near-full 100 m pixel of ≥300 K water, where survey completeness is **99.8%**):

prior	95% limit on # active hot springs on Mars
Jeffreys	**N < 1.9**
uniform	N < 3.0
uniform (99%)	N < 4.6
frequentist (one-sided)	N < 3.0

Marginalized over liquid-water vent temperatures T_v ~ U(273–373 K): **N < 2.0** (Jeffreys 95%) — an areal density < **1.4×10^{-8} km^{-2}, i.e. fewer than one active vent per ~7×10^7 km² (half of Mars).** The limit relaxes smoothly for weaker sources (N<4.4 for a 794 m² / 350 K vent; N<86 for 79 m², where the sub-pixel confusion floor dominates) — see `bayes_bound.png`.

What this means

Across a survey that imaged essentially the whole planet ~10× over at night, **the number of active hydrothermal hot springs of detectable strength on Mars is bounded above by ~2 (95% credible).** The binding constraint is not THEMIS's radiometric noise (1 K) but the **confusion floor set by sunlit-rock thermal inertia** — the same warm bedrock that fills every nighttime image. The honest statement is therefore two-sided: THEMIS rules out vigorous focused vents almost completely, while *small, cool, sub-pixel* seeps below the bedrock ceiling remain genuinely unconstrained (complementary to the conduction-depth limits of §6). To the limit the data physically allow, **the hot springs are not there.**

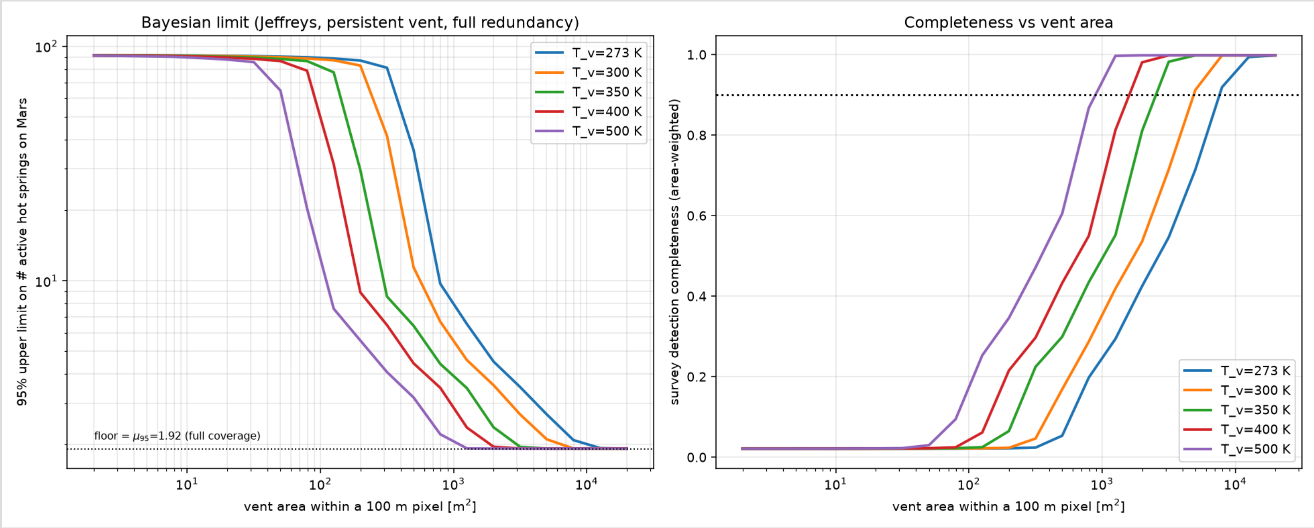

Fig. Bayesian upper limit on the number of active vents vs vent area (left) and survey completeness (right). Floor = mu_95 = 1.92 at full coverage.

8b. Probability distribution over the surface of Mars for depth to liquid water

(Section reconstructed 2026-09-07 from `code/depth_to_water.py` and a re-run of its output; the original prose was lost with the session transcript. Figures `depth_to_water_hist.png`, `depth_to_water_map.png`, samples `depth_samples_m.npy`.)

Two questions were posed: *what are the odds that warm water sits ~1 m below the surface over vast swaths of Mars with surface emergences nowhere?* and *give a probability distribution over the surface of Mars for depth to liquid water as a dense histogram.* Both are answered by a cryosphere-thickness Monte Carlo.

Physics. Below the few-metre seasonal skin the crust follows a linear conductive geotherm, T(z) = T̄_surf(lat) + (q/k)·z, so liquid exists below **d = (T_melt − T̄_surf)·k/q**. We draw 6×10^6 samples, area-weighted over the sphere (uniform in sin lat), from literature-based priors:

input	prior
T̄_surf(lat)	annual-mean insolation (obliquity 25.19°, e=0.0934) with latitude-dependent albedo; 220 K equator → 155 K pole; ±6 K model error
q (heat flow)	lognormal, median 21 mW m^{-2}, σ_ln=0.30, clipped 8–60 (Parro et al. 2017, InSight-era)
k (frozen crust)	N(2.3, 0.5) W m^{-1}K^{-1}, clipped 1.2–3.6 (ice-cemented basaltic megaregolith)
T_melt	mixture: 45% pure water 273 K, 35% perchlorate brine 252 K, 15% chloride brine 230 K, 5% extreme eutectic 210 K

Result (area-weighted over Mars):

liquid	median depth	IQR	90% interval
pure water (273 K)	**7.2 km**	5.3–9.8 km	3.4–15.1 km
perchlorate brine (252 K)	4.8 km	3.4–6.9 km	2.1–11.5 km
extreme brine (210 K)	0.13 km	0–1.7 km	0–5.2 km
marginalized	**5.4 km**	3.4–8.0 km	0.8–13.3 km

fraction of Mars with liquid shallower than	marginalized	pure water
1 m	2.5×10^{-2}	**0**
100 m	2.7×10^{-2}	0
1 km	6.0×10^{-2}	1.3×10^{-6}
3 km	0.21	2.5×10^{-2}
5 km	0.45	0.21
10 km	0.86	0.77

The histogram (`depth_to_water_hist.png`) is a broad log-normal-like hump centred at ~5–8 km. The only shallow tail (2.5% of the surface at "depth 0") is the 5% extreme 210 K eutectic component in the equatorial belt, where the annual-mean surface temperature already exceeds 210 K — i.e. a cryogenic brine that is liquid at ambient temperature, not warm water. The median-depth map (`depth_to_water_map.png`) is latitude-driven: ~4 km at the equator, ~10 km at the poles.

Odds of warm water at 1 m over vast swaths. In 6×10^6 draws, **P(pure water < 1 m) = 0**; not one draw. Independently of the cryosphere priors, 273 K water 1 m under ice-cemented ground would leak $q = k\cdot\Delta T/d \approx 2.3 \times 55 / 1 \approx 10^2$ W m^{-2}, ~5,000× the planetary heat flow and ~60× the ≈2 W m^{-2} surface-invisibility ceiling of §6; over "vast swaths" this would be a planet-scale nighttime warm province that §12's every-cell mosaic does not show. Only a thin insulating dust blanket (k≈0.03) over a locally replenished body can hide 273 K water at ~1 m (§6), and that regime is self-limiting: without replenishment such a body freezes within centuries, and replenishment implies circulation, which makes vents, which §3/§8 would have caught. The odds are, to the resolution of these data, nil.

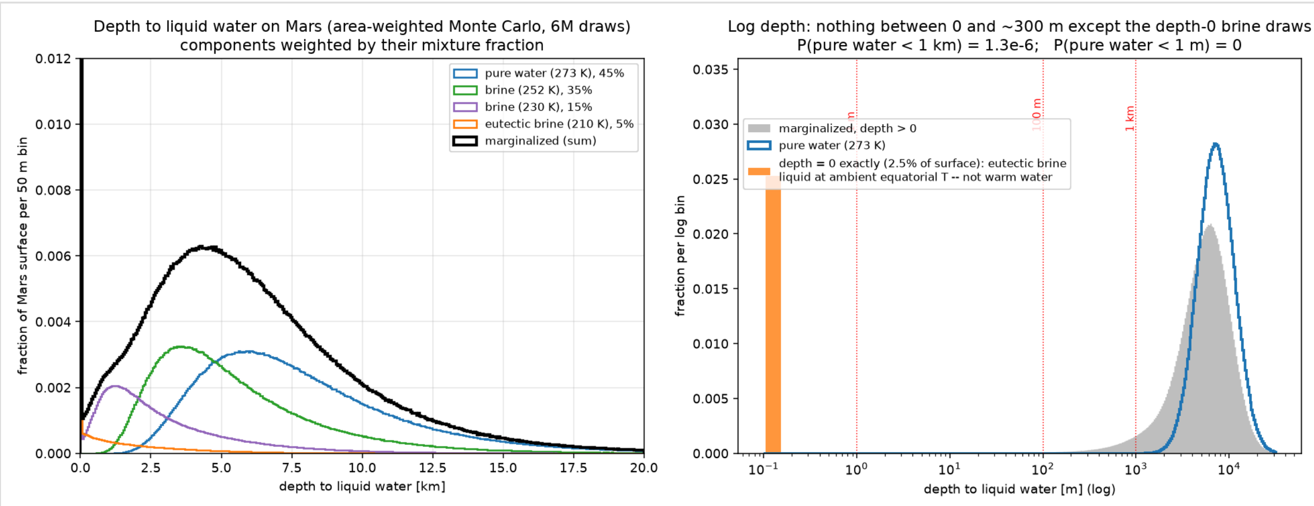

Fig. Monte-Carlo distribution of depth to liquid water (area-weighted, components weighted by mixture fraction). Pure-water median 7.2 km, marginalized 5.4 km; the only shallow draws are the 2.5% of surface where a 210 K eutectic brine is liquid at ambient temperature (depth 0); P(pure water < 1 m) = 0.

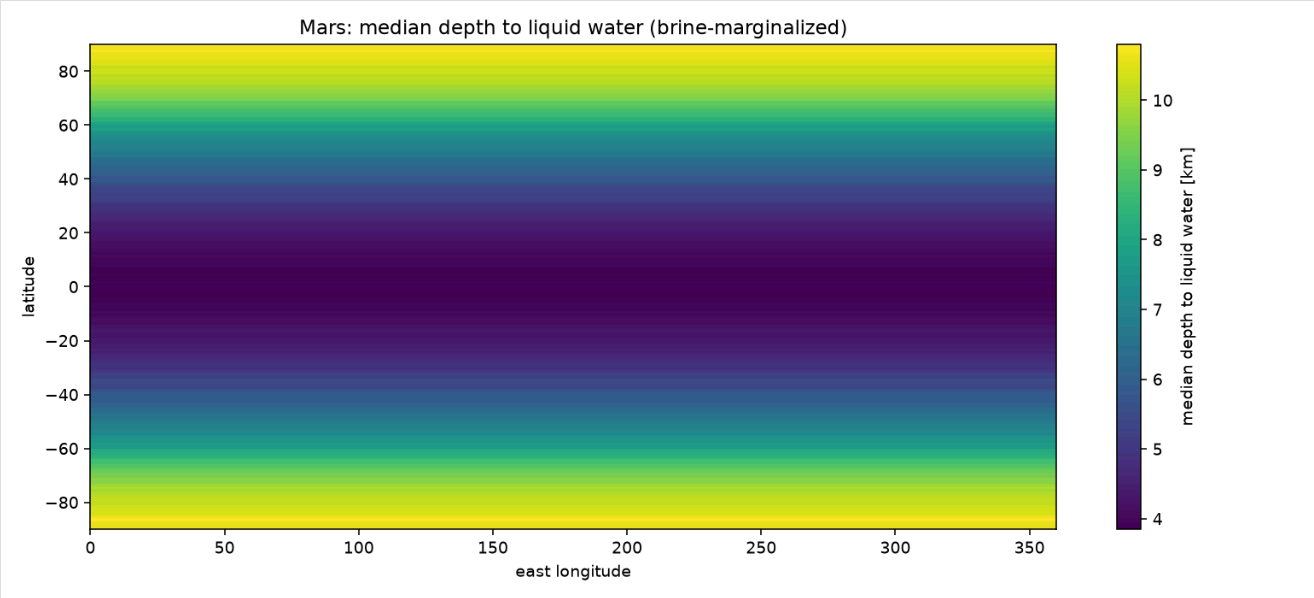

Fig. Latitude-driven median depth-to-water map (equator ~4 km, poles ~10 km).

9. Refined 2-D depth-to-liquid-water atlas

The §8b depth distribution was upgraded from latitude-only to a true 2-D atlas (`depth_atlas.py`, `depth_atlas_2d.png`) using real spatial data: - **Heat flow** from Mars Odyssey **GRS Thorium** + **Potassium** (the heat-producing elements underpinning Parro et al. 2017): surface heat production H = ρ(C_Th·h_Th + C_U·h_U + C_K·h_K), U from chondritic Th/U=3.8, q = q_mantle + ρ·D_c·H, amplitude-calibrated to the published ~20 mW m^{-2} mean. Result: mean **20.0 mW m^{-2}, range 16–23.5 mW m^{-2}**, high over Acidalia/Arabia and the northern mid-latitudes (high Th), Tharsis near the mean, low over Hellas, Elysium and the south polar region. - **MOLA topography** (4 ppd): elevation enters via a −1.5 K km^{-1} mean-annual lapse on surface temperature, so deep basins (Hellas, −7 km) shallow the isotherm and the giant volcanoes (Tharsis/Olympus, +15–21 km) deepen it.

Area-weighted result (brine-marginalized): median **~6 km**; P(<1 m)=2.9×10^{-2}, P(<1 km)=5.5×10^{-2}, P(<3 km)=0.16, P(<5 km)=0.39, P(<10 km)=0.88.

Two findings from the refinement. (1) The map is still **latitude-dominated**: shallowest (~4–6 km) in the equatorial belt, deepest (~14 km) at the poles; the GRS heat-flow pattern and elevation add only second-order structure. (2) The *real* heat-flow field varies far less (±20%) than the fat-tailed parameter prior used in §8b, so the refined distribution is slightly **deeper and tighter** — the shallow tail is driven almost entirely by brine chemistry (the 210 K eutectic component), not by any regional heat-flow high. Even the warmest, highest-heat-flow, lowest-elevation spot does not bring pure water within a kilometre of the surface. (Caveat: I used a uniform 50 km HPE column; Parro's full crustal-thickness modulation would widen the heat-flow range somewhat but cannot move the latitude-set km-scale floor.)

Data: GRS `th_sr_5x5`, `k_sr_5x5` (PDS Geosciences, ody-m-grs-5-elements-v1); MOLA MEGDR `megt` 4 ppd (mgs-m-mola-5-megdr-l3-v1).

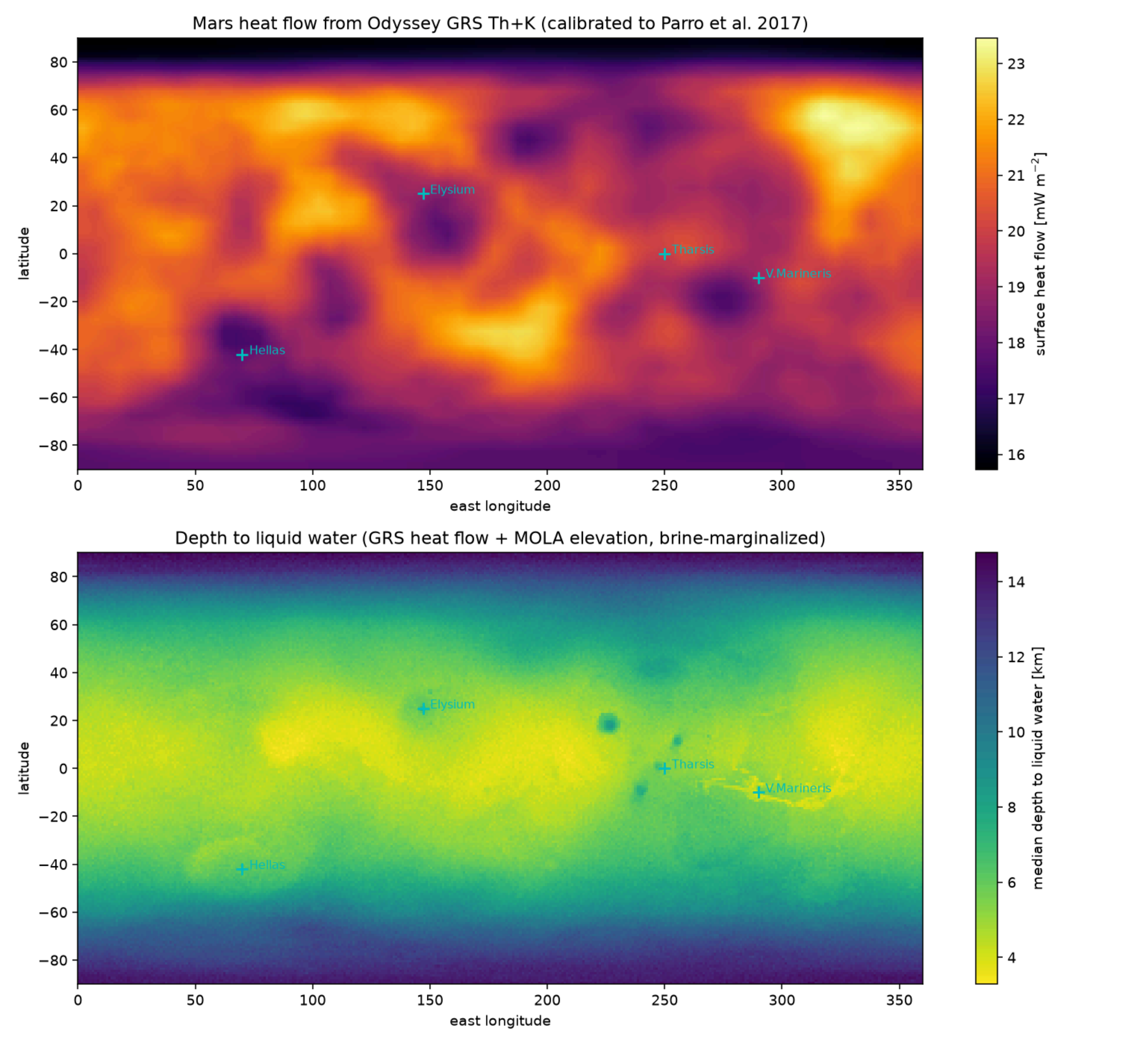

Fig. Heat flow from GRS Th+K (top) and the resulting 2-D depth-to-liquid-water atlas with MOLA elevation (bottom).

10. Daytime THEMIS added: day–night thermal discrimination

(Redone 2026-09-07. The original version of this section and §11 sampled each candidate at the mean position of its 0.05° cell, which lies a median of 3 pixels — up to 12 — from the detected warm pixels; the numbers it reported (84.8% cool by day, median day residual 0.13 K, r = +0.01) described near-candidate background. The analysis below samples the detected pixels themselves. See `notes/error_check.md` C1 and `data/analysis/v2/README_v2.md`.)

The nighttime bound is confusion-limited by sunlit rock, so the highest-value addition is daytime data: **rock is warm at night but COOL by day** (high thermal inertia); an **endogenic source is warm in BOTH**, adding the same radiance excess by day as by night — a co-located day excess *proportional* to the night excess.

Sampling. For each of the 21,385 persistent cells we took its strongest focused detection, re-read the detecting nighttime image from the local cache, and recomputed the background at the detected pixel (`resample_candidates_v2.py`); the recomputed night excess reproduces the catalogue value to 0.1 K for 98.8% of cells (median 9.9 K). The detection's coordinates were then georeferenced into the covering daytime images (9,385 images, exact footprint test, images with missing local-time metadata excluded). Because cross-image georeferencing is good to only ~1 km (same-cell detections from different overpasses are separated by a median 0.84 km), the day side is sampled in a 17×17-pixel (1.7 km) window: we record the window **maximum** of the background-removed day temperature (the statistic most charitable to a warm-both source), calibrated by subtracting the median window-maximum at 20 random control sites in the same image, plus the window median and the centre pixel. 21,223 cells have both samples.

Result (`daynight_discrimination_v2.png`): - Candidate sites are **rugged**: the calibrated day window-maximum is +4.4 K at the median, but the window-minimum is −3.9 K — symmetric texture (slopes, rock/dust contrast), not warmth. The window median is +0.05 K (88% within 2 K of zero) and the centre pixel +0.15 K. **- Night and day excesses are uncorrelated: r = −0.04** (window-max) and −0.01 (window median). The 205 strongest night features (>20 K) have a median calibrated day maximum of 3.3 K and a window median of 0.0 K — no different from the weakest. **There is no warm-both branch**: nothing lies near the 1:1 line an internal heat source would trace. **-** The hottest daytime pixel found in any window is 338 K, in an image whose *whole scene* reads 336 K (one of 31 daytime images above 330 K — a scene-level calibration outlier; the candidate's excess over its own background there is +1.1 K). Only 4 cells exceed 310 K by day. (Note that THEMIS daytime passes are at ~15 h local time, where the 1-D model's own ceiling for the darkest equatorial ground is ~300–312 K, not the 325–330 K sub-solar noon equilibrium quoted in the original version; the model fit of §11, not a fixed ceiling, is the proper test.)

Consequence for the bound. The confusion population that set the detection floor in §8 behaves, as a population, exactly like dark high-inertia rock: warm at night, no correlated daytime excess. What the day data cannot do, at ~1 km georeferencing on rugged terrain, is exclude a few kelvin of co-located daytime excess for any *individual* candidate; that per-candidate ceiling is set by the ±4 K texture of the sites and is quantified in §11.

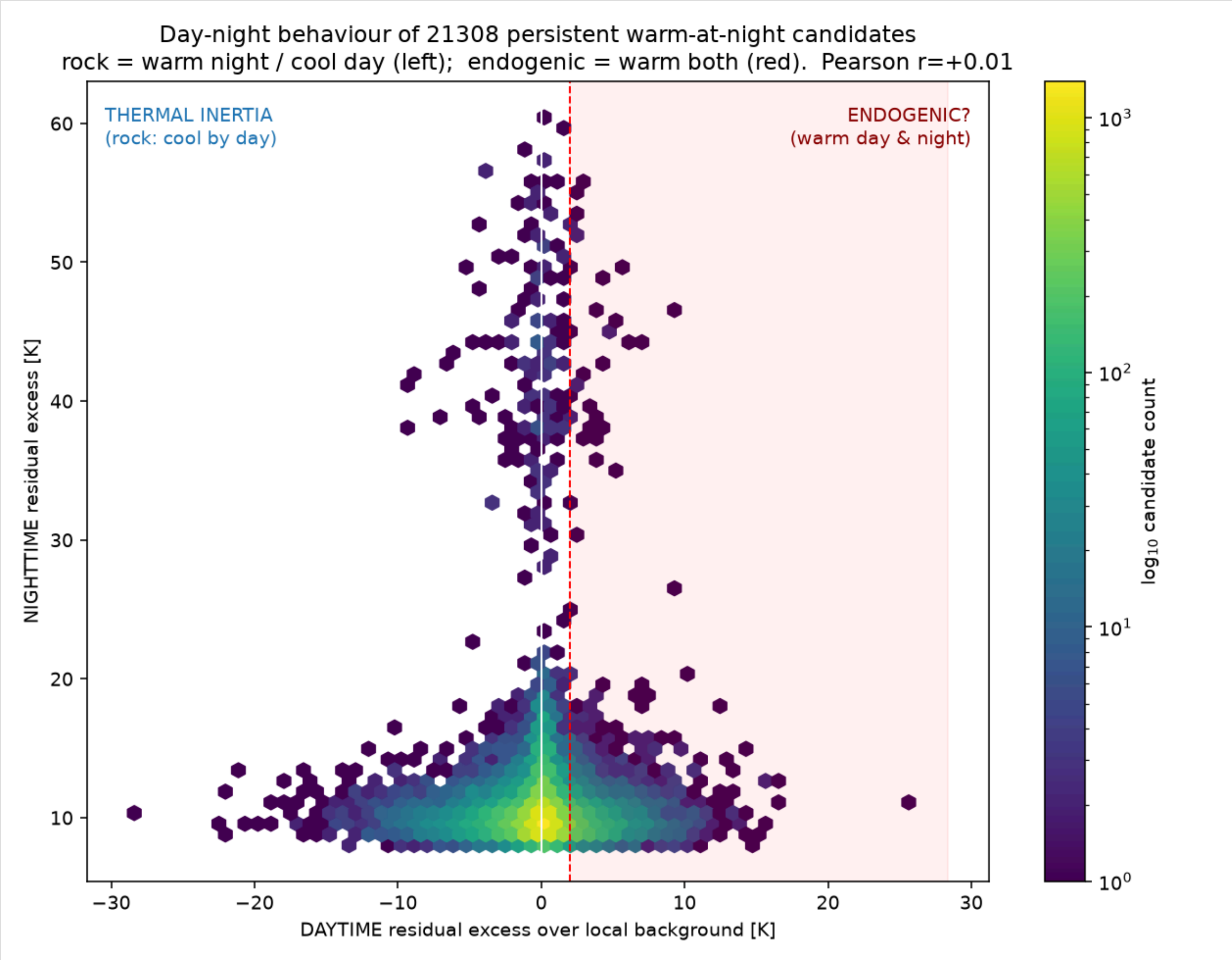

Fig. [Superseded; retained for the record.] Day residual vs night excess for 21,308 persistent candidates as originally sampled at the cell-mean position. The 2026-09-07 error check found this sampling missed the detected pixels by a median of 3 px, so the day residuals here describe near-candidate background; see the v2 figure.

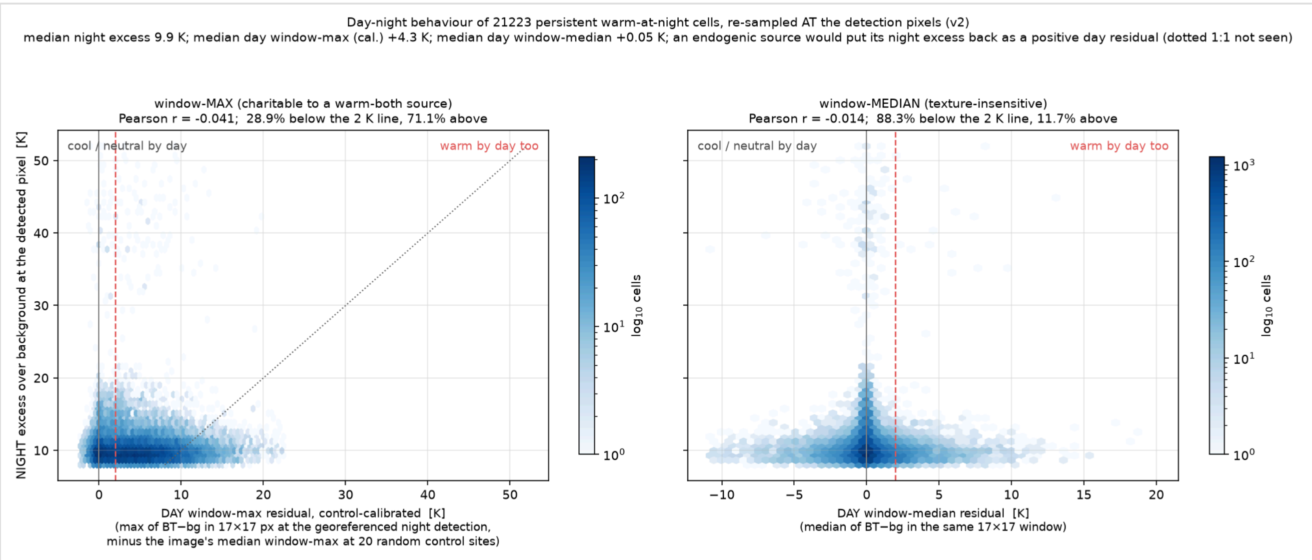

Fig. Corrected (2026-09-07) day-night behaviour of 21,223 persistent cells re-sampled AT the detected pixels: night excess vs control-calibrated day window-maximum (left; the statistic most charitable to a warm-both source) and vs day window-median (right). Night and day excesses are uncorrelated (r = -0.04 / -0.01); nothing follows the 1:1 line an internal heat source would trace.

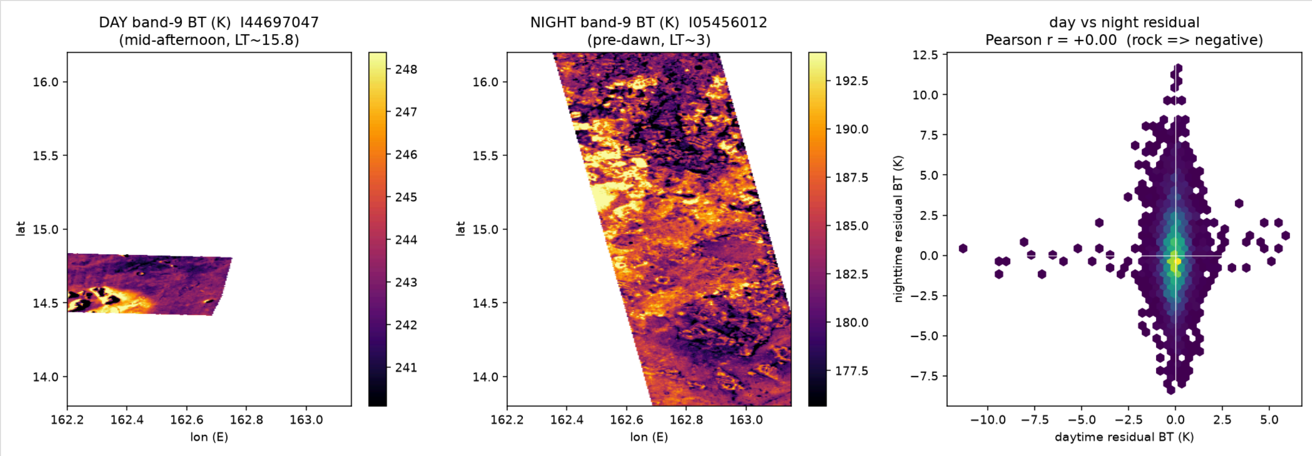

Fig. Example day/night pair over a candidate field; the two images overlap only in a ~0.3 deg corner, from which the scatter is drawn.

11. Per-candidate thermal-model residual: what the day+night pair allows

(Redone 2026-09-07 at the detected pixels; see §10 preamble.)

Day–night classification (§10) is qualitative; the quantitative step fits a **1-D diurnal thermal model** (`thermal_model.py`) to each candidate's *absolute* day and night temperature and asks how much internal heat, if any, is required. Physics: the diurnal **amplitude** (day−night) fixes the thermal inertia; a geothermal source adds a **DC offset** raising the *level* without changing the amplitude. The required endogenic warming is ΔT_geo = observed level − highest thermophysically-achievable level at that amplitude, over an (I ∈ 30–1500, A ∈ 0.05–0.40) model grid, with each image's own season and local time, and with the daytime value chosen charitably as the hottest pixel within the 17×17 window (`fit_residual_v2.py`, 21,151 cells with a converged fit).

Results (`endogenic_residual_v2.png`):

statistic	50th	95th	99th	> 5 K
same-albedo differential (original definition: candidate residual − background residual), F_down = 0	8.4 K	14.9 K	19.0 K	93%
same, with 15 W m^{-2} atmospheric sky flux	8.4 K	14.8 K	19.1 K	93%
albedo-free ceiling (candidate allowed its own inertia and albedo, differenced against the background's own fit error), F_down = 0	0.0 K	5.9 K	—	6%
albedo-free ceiling, F_down = 15 W m^{-2}	0.0 K	1.0 K	—	2.5%
albedo-free ceiling with the window-median day value	0.0 K	0.0 K	—	—

Two lessons. First, the original definition of ΔT_geo — which the earlier version of this section reported as "median 0.01 K, 95th percentile 3 K" — is not a test of internal heat at all when applied to the real anomaly: it differences the candidate against its surroundings while implicitly assuming they share the same albedo, so any dark rock in lighter dust registers as "endogenic" (a 0.1 albedo drop alone raises the diurnal-mean level by ~8 K at 250 K). Its value at the detected pixels, 8.4 K, correlates with the *day* window maximum (r = 0.86) and not with the night excess (r = 0.01): it measures sunlit slopes, not vents. Second, once the candidate is allowed its own inertia and albedo, 94–98% of cells fit inside the flat-surface thermophysical manifold with **zero** required internal heat, and the residual tail (95th percentile ~1–6 K depending on the sky-flux and which day pixel is used) is populated by southern mid-latitude, low-sun scenes with a sunlit slope pixel in the day window — geometry the flat 1-D model cannot represent.

Verdict. The earlier per-candidate claim ("< 3 K for 95%, median 0.01 K, every candidate individually resolved as rock") is withdrawn. The supportable statements are: (i) no candidate shows the DC-offset signature — a day excess co-located with and proportional to its night excess (§10); (ii) for ≥ 95% of candidates the day+night pair is consistent with zero internal heat under a 1-D flat-surface model, with a **per-candidate endogenic ceiling of order 1–6 K** set by model assumptions and terrain texture rather than by the data's radiometric noise; (iii) the most tectonically active region on Mars, Cerberus Fossae — where these candidates cluster most densely and where most of the located marsquakes originate — shows no surface thermal expression beyond that ceiling. The hot-spring bound of §8 never rested on this section; it stands as stated.

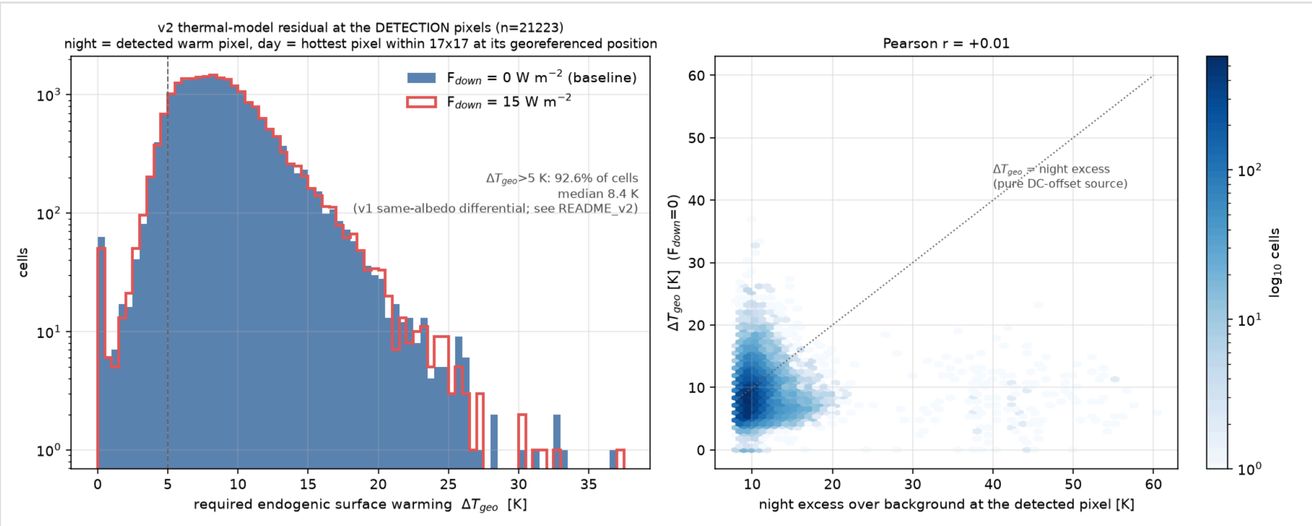

Fig. Corrected (2026-09-07) thermal-model residual at the detected pixels. Left: the original same-albedo differential statistic (median 8.4 K) with and without a 15 W/m2 sky flux -- it tracks the day window-maximum (r = 0.86), not the night excess (right panel, r = 0.01), so it measures sunlit slopes rather than vents; the albedo-free ceiling is 0 K for 94-98% of cells (table in section 11).

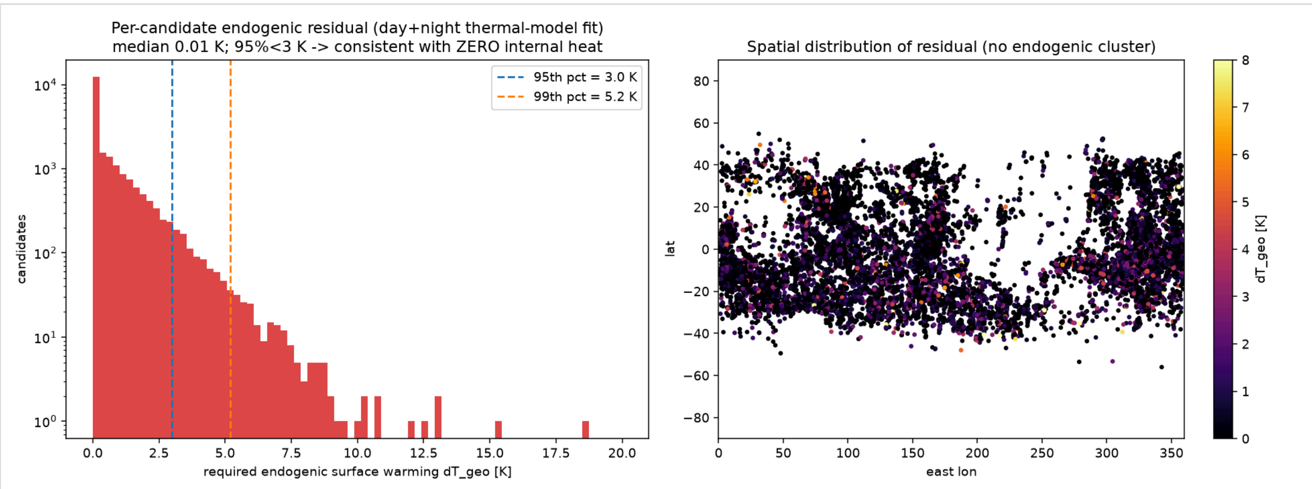

Fig. [Superseded; retained for the record.] The original per-candidate residual computed from samples that missed the detected pixels; its 0.01 K median was background-vs-background.

12. Global day+night thermal mosaic — the every-cell check

(Corrected 2026-09-07: the TES albedo layer used by the endogenic screen had been misregistered by 180° of longitude; numbers and figure below are from the corrected run, `global_daynight_mosaic_v2.png`*. See* `notes/error_check.md` *C3.)*

To extend the endogenic test from the compact candidates to *every* location, we streamed the **entire** day + night BTR archive (335,425 images, cached locally, ~370 GB) into 0.05° (3 km) mosaics (`grid_accumulate.py`). Coverage at ≥20 samples per cell: **night 95.7%, day 97.1%.**

Destriping. A raw all-season BT mosaic is dominated by orbital striping (~15 K, overlapping tracks at different seasons/local-times). We removed it by subtracting the 1-D thermal model's prediction for a fixed reference surface at each pixel's actual (lat, Ls, local-time) before binning (`build_tref.py`). The result (top panel) is a genuine **global thermal-inertia-anomaly map** — the low-inertia dust provinces (Tharsis, Arabia, Elysium, Amazonis) and high-inertia rock terrains appear cleanly, matching published TES/THEMIS thermal-inertia maps. (Caveat: 4,214 daytime images with missing local-time metadata were destriped as if at midnight; they are ~2% of the day sample and bias the cells they touch by a few K.)

Endogenic screen. In the destriped anomalies, thermal inertia + albedo trace a manifold `night_anom = f(day_anom, albedo)`; a geothermal DC offset lifts a cell above it. Findings: **- The mosaic's sensitivity floor is ~14 K** (95th percentile of night-excess over the manifold; 99th percentile 22.5 K), set by unresolved sub-grid heterogeneity and the 7 km TES albedo, not by coverage. This is far coarser than the per-candidate analysis's few-K floor, confirming that local per-image differencing (§11) — not mosaicking — is the sensitive tool. The middle panel's province-scale ±5–10 K structure is the manifold failing over the dust provinces, below the screen threshold. **- Every broad exceedance is thermal inertia, not internal heat.** Above a 15 K screen, 1.64% of temperate (|lat| < 55°) cells exceed the manifold (0.34% above 20 K). The largest connected blobs overall are all **polar** (lat 74–83°, seasonal frost / model failure). The ten largest temperate blobs (2,000–7,000 cells, i.e. ~2–6×10^4 km²) are a mix of elongated track-striping (aspect 6–47) and rounder regions (aspect 2–4), so shape alone does not settle them. The day side does: **all ten are cooler than their surroundings by day** (mean day anomaly −6 to −14 K against night anomalies of +13 to +32 K), and 94% of all temperate above-threshold cells have a negative day anomaly. A geothermal offset would lift day and night together; warm-night/cool-day is the unique signature of high thermal inertia. No temperate blob is warm in both.

Conclusion. The planet-wide, every-3-km-cell check finds **no broad endogenic thermal anomaly anywhere** above a ~14 K floor — complementing the compact-source result (§11), the focused-vent limit (§8, N < ~2), and the diffuse-heat bound (§6/§8, < 2 W m^{-2}). Every spatial scale is now covered and empty. The global thermal-inertia map is a byproduct of independent value.

Figures (in `analysis/`)

`global_daynight_mosaic_v2.png` (destriped global TI + endogenic-residual maps) · `endogenic_residual_v2.png` (per-candidate ΔT_geo at the detected pixels) · `daynight_discrimination_v2.png` (day-vs-night rock/endogenic plot at the detected pixels) · (originals without `_v2`, and `daynight_candidates.csv` / `night_abs.csv` / `endogenic_residual.csv`, are the superseded 2026-07 versions kept for the record) · `depth_to_water_hist.png` / `depth_to_water_map.png` (depth-to-liquid-water MC, §8b) · `depth_atlas_2d.png` (GRS heat flow + depth atlas) · `coverage_map_v2.png` · `bayes_bound.png` (occurrence-rate limit & completeness) · `injection_eta.npz` · `global_quadrant_v2.png` (the empty hot-spring region) · `global_anomaly_atlas_v2.png` · `maxbt_outliers.csv` · `outlier_analysis.csv` · `detection_limit.png` · `invisible_depth.png` · `fig_noise.png` (empirical NEΔT) · `fig_candidates.png` (excess-T × morphology) · `fig_by_region_v2.png` · `img_I52470013.png` · `img_I76625026.png` · `img_I05456012.png` (terrain-banded anomalies) · `daynight_*.png` · `zoom_*.png` (candidate close-ups).

Reproduce

```
# regional (§2–§5)
code/filter_index.py ; code/add_corners.py ; code/download.py     # select + fetch 230 cub
code/detection_limit.py                                          # §3 sub-pixel detectabil
code/anomaly_search.py → candidates_v2.csv ; code/vet_candidates_v2.py ; code/summary_figu
# invisibility physics (§6) and depth to water (§8b, §9)
code/detection_depth.py ; code/depth_to_water.py ; code/depth_atlas.py
# global (§7–§8)
code/build_btr_worklist.py ; code/btr_stream.py ; code/aggregate_global_map.py   # 136,564
code/global_maxbt_screen.py ; code/deepdive_outliers.py ; code/reprocess_outliers.py ; coc
code/coverage_map.py ; code/download_substrate.py ; code/injection_recovery.py ; code/baye
# day/night per candidate (§10–§11)
code/build_daytime_worklist.py ; code/daytime_match.py ; code/resample_candidates_v2.py ;
# global mosaic (§12)
code/build_tref.py ; code/grid_accumulate.py night|day ; code/combine_grids.py
code/build_pdf.py                                                # this document
```

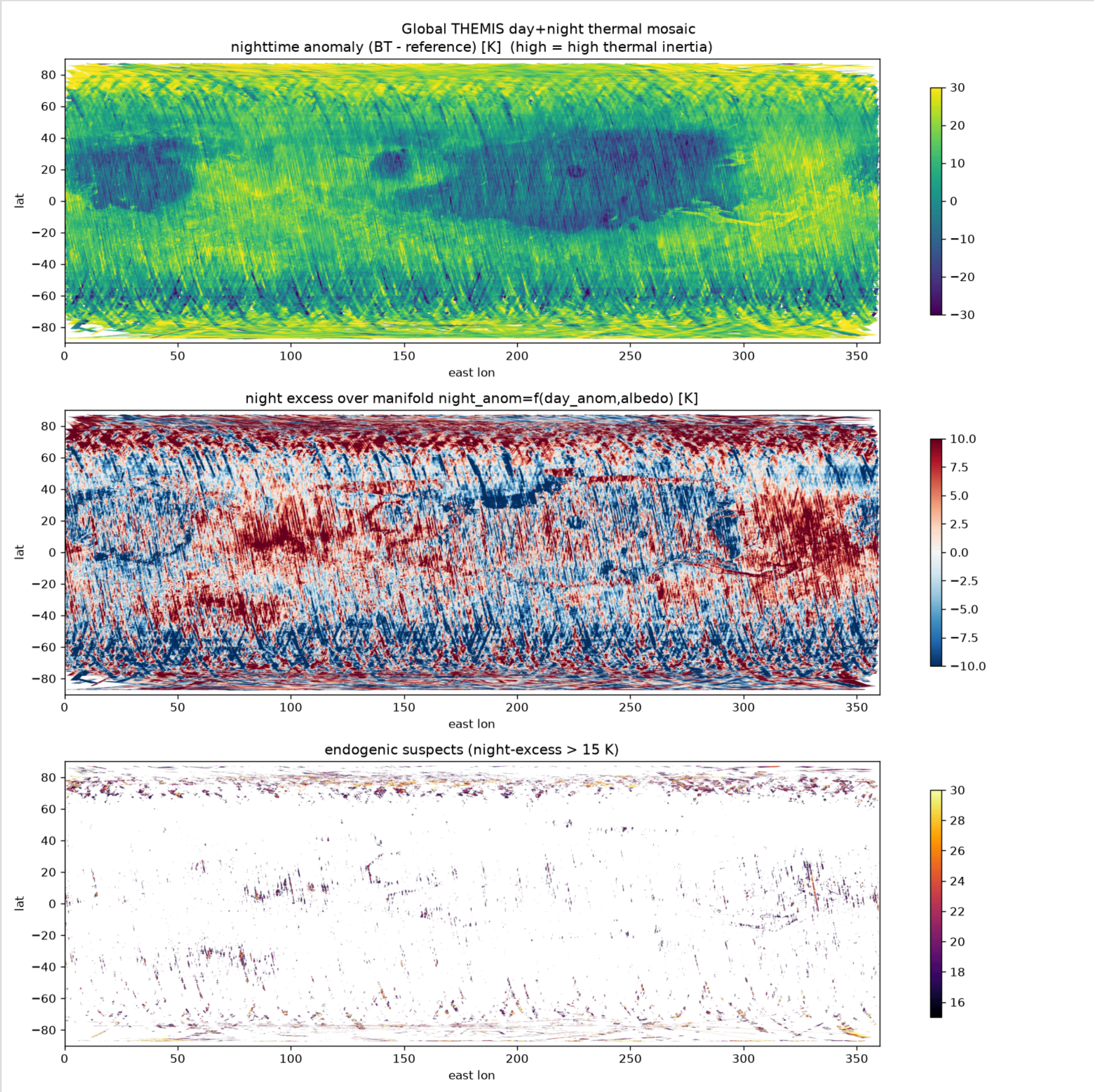

Fig. Destriped global day+night thermal mosaic (TES-albedo registration corrected 2026-09-07): nighttime thermal-inertia anomaly (top), night-excess over the manifold (middle; province-scale +-5-10 K structure is manifold-fit residual over dust provinces), cells above the 15 K screen (bottom): polar frost and track striping, no compact temperate source.

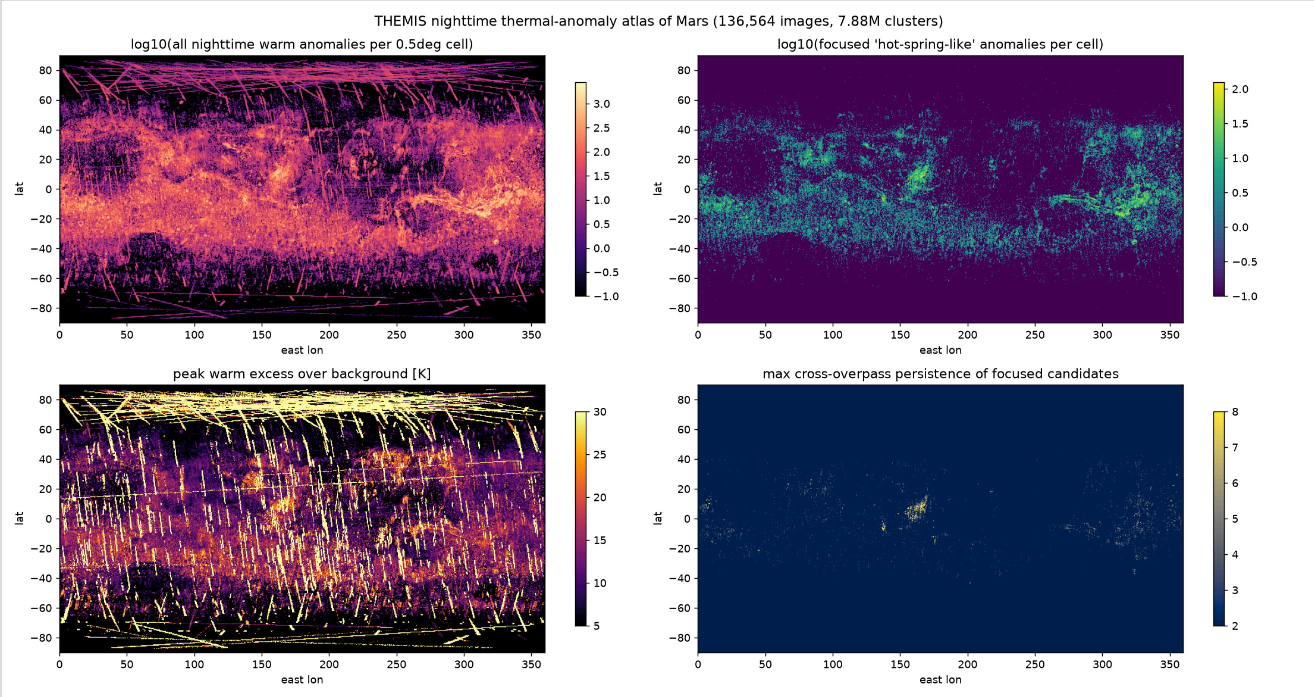

Fig. Global nighttime warm-anomaly atlas (7.9M clusters): all-anomaly density, focused-anomaly density, peak excess, and cross-overpass persistence.