

# BurnWatch

Satellite Detection of Open Burn Sites Using Sentinel-2 Change Analysis

*Technical Replication Guide — Deploying the System for Any County*

Version 1.0 • July 2026 • Reference deployment: Clinton County, Missouri

# Contents

Contents.....	2
1. Purpose and Concept.....	3
2. System Architecture.....	3
3. Data Sources and Requirements.....	4
3.1 Sentinel-2 imagery (free, global) .....	4
3.2 County parcel layer (the one local requirement).....	4
4. Detection Methodology.....	4
4.1 The Normalized Burn Ratio (NBR).....	4
4.2 Cloud masking and the rolling baseline .....	5
4.3 Burn candidate criteria .....	5
4.4 Clustering, merging, and the size band-pass .....	5
4.5 Persistence confirmation (two-look verification) .....	5
4.6 Parcel attribution, site recurrence, and scoring .....	6
5. Calibration — Method and Reference Results .....	6
6. Adapting the System to a New County .....	7
7. Installation and First Run .....	7
8. Automated Weekly Operation.....	8
9. Outputs and Data Model .....	8
10. Established-Pit Detection: the Three-Stage Architecture.....	9
11. Validation Against Ground Truth .....	9
12. Limitations and Responsible Use .....	11
Appendix A. Calibrated Configuration Reference (Clinton County, MO).....	11

## 1. Purpose and Concept

BurnWatch is a self-hosted monitoring pipeline that detects new open-burn sites across an entire county using free European Space Agency (ESA) Sentinel-2 satellite imagery, joins each detection to the county parcel layer to identify the property owner, and ranks the results into an investigation queue. Its purpose is to surface candidate sites of unlawful trade-waste burning — for example, commercial tree services disposing of customer debris in open pits — so that limited investigative effort (drive-by verification, regulatory complaints) is spent only on the highest-probability targets.

The core principle is spectral change detection. Burned ground has a distinctive signature in shortwave-infrared imagery. By comparing every new satellite pass against a rolling per-pixel baseline, the system flags exactly what changed, when it changed, and on whose land — county-wide, roughly twice per week, at zero data cost.

**What a detection is:** evidence that a burn-scale spectral change occurred at specific coordinates within a specific date window. **What it is not:** proof of who burned, what material was burned, or whether the burn was lawful. Detections are targeting leads that require ground verification before any public accusation or enforcement referral. See Section 10.

## 2. System Architecture

The pipeline is roughly 1,500 lines of Python across six modules, with no server infrastructure required. It can run on a desktop, a scheduled cloud runner (GitHub Actions), or any Linux host.

Module	Responsibility
config.py	Every tunable in one file: credentials, county bounding box, detection thresholds, size band, scoring weights, parcel field mappings.
sentinel.py	Copernicus Data Space API client: OAuth token management, catalog search (which dates have new scenes), and Process API retrieval of the three needed bands, decoded via tiff file.
detect.py	The science: NBR computation, cloud masking, dNBR change detection, cluster merging and consolidation, size/shape filtering, parcel spatial join (shapely STRtree), and priority scoring.
store.py	SQLite persistence: processed-scene ledger, site registry with recurrence counts, detection records, and per-tile baseline arrays saved as NumPy files.
run.py	Orchestrator: finds unprocessed scene dates, runs detection per tile, updates state, exports outputs. Supports --backfill, --export-only, --test-auth.
make_review_map.py	Renders a fully self-contained HTML triage console (Leaflet + satellite basemap + ranked queue) with the detection data embedded inline — opens from disk, no web server.
watch.py	Watchlist: maximum-sensitivity absolute-NBR time series over known sites (dark-core tracking of the darkest decile). Immune to the baseline blind spot; produces per-site dated burn-event records.
scar_sweep.py	Multi-signal established-pit sweep: never-greens (max NDVI), persistent non-vegetation (mean NBR), pit-scale size band, elongation-aware, activity-ranked.
fuse.py	Evidence fusion: merges sweep candidates with confirmed burn events and watchlist events into one convergence-ranked list.

Module	Responsibility
vision_review.py / reject.py	Vision stage: classifies top fused candidates on MSDIS 6-inch aerial chips via the Claude vision API; reject.py records human review verdicts and rebuilds the census.

**Data flow per run:** catalog search → fetch B08/B12/SCL per tile per new date → cloud mask → NBR → dNBR vs. baseline → threshold → cluster/merge/filter → parcel join → site match & score → SQLite → GeoJSON + HTML export → baseline roll-forward.

### 3. Data Sources and Requirements

#### 3.1 Sentinel-2 imagery (free, global)

Sentinel-2 is a pair of ESA satellites imaging all global land every ~5 days combined. Level-2A products (atmospherically corrected surface reflectance) publish within hours of acquisition on the Copernicus Data Space Ecosystem ([dataspace.copernicus.eu](https://dataspace.copernicus.eu)). The pipeline uses three layers per scene:

- B08 — near-infrared (NIR), 10 m native resolution.
- B12 — shortwave infrared (SWIR), 20 m native, upsampled to 10 m by the API. This band carries the burn signature.
- SCL — ESA's per-pixel scene classification, used to mask clouds, shadows, water, and snow.

Access requires a free Copernicus account and an OAuth client (created in the Sentinel Hub dashboard at [shapps.dataspace.copernicus.eu/dashboard](https://shapps.dataspace.copernicus.eu/dashboard)). Free accounts include a monthly processing quota; a single-county weekly workload consumes a negligible fraction of it.

#### 3.2 County parcel layer (the one local requirement)

The only county-specific data requirement is a parcel polygon layer with ownership, as GeoJSON in WGS84 (EPSG:4326). Sources vary by county: an open data portal, the assessor's GIS vendor, a Sunshine Law / public records request, or export from an existing GIS viewer. The reference deployment used 13,330 parcels (~8.5 MB GeoJSON). Three attribute fields are required, and their names are mapped in config.py:

```

PARCELS_PATH      = "parcels.geojson"
PARCEL_ID_FIELD   = "parcel_num"      # unique parcel identifier
PARCEL_OWNER_FIELD = "owner"          # deeded owner name
PARCEL_ACRES_FIELD = "deeded_acr"     # parcel acreage

```

Geometry may be Polygon or MultiPolygon; invalid geometries are automatically repaired (zero-width buffer) at load time.

### 4. Detection Methodology

#### 4.1 The Normalized Burn Ratio (NBR)

For every clear pixel the system computes  $NBR = (B08 - B12) / (B08 + B12)$ . Healthy vegetation reflects strongly in NIR and weakly in SWIR, giving NBR values around +0.3 to +0.7. Combustion destroys that structure: char and ash absorb NIR and reflect SWIR, driving NBR to near or below zero. The differenced

index,  $dNBR = NBR_{baseline} - NBR_{current}$ , isolates change: large positive  $dNBR$  means the surface lost vegetation signal since the last look.

## 4.2 Cloud masking and the rolling baseline

Pixels are trusted only when the SCL classification marks them as vegetation (4), bare ground (5), or unclassified (7); clouds, shadows, water, cirrus, and snow are excluded. A scene tile is skipped entirely below 25% clear ( $MIN\_CLEAR\_FRACTION$ ). Each tile maintains a rolling baseline: the most recent clear observation of every pixel. After each processed scene, all clear pixels — including newly burned ones — are folded into the baseline. This is deliberate: a scar flags once, when fresh, rather than re-flagging on every subsequent pass. “Recurrence” therefore means new burn events, not the same scar seen twice.

## 4.3 Burn candidate criteria

A pixel is a burn candidate only when all conditions hold (calibrated values, Section 5):

- Clear in both the baseline and the current scene;
- $dNBR \geq 0.32$  — substantial loss of vegetation signal; and
- $current\ NBR \leq 0.0$  — the surface now looks like char/ash, not merely bare dirt. This is the physical discriminator that rejects tillage.

## 4.4 Clustering, merging, and the size band-pass

Candidate pixels are grouped into objects with an 8-connected labeling after a 3-pixel binary dilation ( $MERGE\_RADIUS\_PX$ ), which fuses patchily-detected fragments of one burn into a single object before any size test. Objects then pass a two-sided size filter measured in true burned pixels:

- Floor — 0.08 acres ( $MIN\_BURN\_ACRES$ ): eliminates single-pixel speckle and trivial backyard piles.
- Ceiling — 5.0 acres ( $MAX\_BURN\_ACRES$ ): discards whole-field agricultural burns, which are typically 10–100+ acres. A commercial disposal pit complex rarely exceeds ~2–3 acres of burn signature.

Two additional guards close known evasion paths of the ceiling:

- Compactness — Polsby-Popper ( $4\pi A/P^2$ )  $\geq 0.12$  rejects linear artifacts: road and stream edges, fence-line burns.
- Group consolidation — same-scene clusters within 250 m of one another are chained (single-linkage), and if the group's total burned acreage exceeds the ceiling, the entire group is discarded. This prevents a large patchy field burn from fragmenting into many individually-passing pieces — a failure mode observed in real calibration data (Section 5).

## 4.5 Persistence confirmation (two-look verification)

Every detection begins life as PROVISIONAL. On the next clear pass over the same ground, the system re-samples the detection's exact pixel footprint: if mean NBR remains at or below  $CONFIRM\_NBR\_MAX$  (0.15) — a scar, which persists for weeks — the detection is CONFIRMED, scored, and counted toward site recurrence. If the surface has recovered to normal reflectance, the detection is RETRACTED as a one-scene artifact. This mechanism eliminates an entire class of failure regardless of cause — unmasked cloud shadow, haze, thin smoke, orbit-seam effects — discovered in calibration when a single scene date (June 1) produced ~900 phantom detections that passed every spectral test but vanished by the following pass. Confirmation adds

one clear-pass of latency (typically 5–14 days), which is immaterial at a weekly monitoring cadence; provisional detections are shown dashed/gray on the review map pending their second look.

#### 4.6 Parcel attribution, site recurrence, and scoring

Each surviving detection polygon is spatially joined against the parcel index (STRtree); the parcel with the greatest overlap is primary, with all intersecting parcels retained. Detections within 250 m (SITE\_MATCH\_RADIUS\_M) of a known site on different dates count as recurrence at that site; same-date sibling fragments never inflate the count. Each detection receives an additive investigation-priority score:

Signal	Weight	Rationale
Base (every detection)	+1.0	Everything that survives filtering merits a look.
Each prior burn date at the site	+2.0	Repeated burning at one location is the signature of an operation, not a one-off legal burn.
Owner matches hot keywords (TREE, STUMP, LOGGING, TIMBER, CLEARING...)	+3.0	Direct trade-relevance to wood-waste disposal.
Owner matches commercial keywords (LLC, INC, SERVICE, EXCAVAT...)	+1.0	Commercial entities are likelier trade-waste generators.
Parcel under 40 acres	+1.0	Small non-agricultural parcels have no on-site source for large repeated burns — material is being hauled in.
Mean dNBR $\geq$ 0.44	+1.0	Moderate-high severity indicates intense, sustained fire.

All weights and keyword lists are configuration values intended to be adapted to local land use and the specific trade being monitored.

### 5. Calibration — Method and Reference Results

Thresholds must be calibrated against a real historical backfill, because the dominant false-positive sources are local and seasonal. The procedure used for the reference deployment, recommended for any new county:

1. Run a 3-month historical backfill at permissive settings (dNBR  $\geq$  0.27, post-NBR  $\leq$  0.10).
2. Review the flag volume and composition per scene date. Identify systematic false-positive patterns rather than judging individual flags.
3. Adjust the smallest number of parameters that removes each pattern; wipe state; re-run the backfill. Repeat until weekly volume is reviewable (single digits to low tens county-wide outside peak season).

#### Reference findings (Clinton County, MO, April–July 2026 backfill):

- Fragmented field burns defeated the acreage ceiling: one ~8-acre pasture burn detected patchily as ~15 sub-ceiling fragments. Remedied structurally with pre-sizing dilation merge plus group-total consolidation (Section 4.4) — not with thresholds.

- Recurrence scoring initially counted same-date sibling fragments as prior events, inflating scores from 1 to 17 within a single scene. Remedied by counting distinct prior dates only.
- Hay cutting and wheat harvest produced the largest false-positive spikes (1,300+ flags on single scene dates matching the regional crop calendar). Cut and harvested vegetation drops NBR like a burn but remains BRIGHT in near-infrared, while char absorbs NIR. Remedied with the `POST_NIR_MAX = 0.18` reflectance ceiling — a second physical discriminator, not a threshold tweak.
- A single scene date produced ~900 artifact detections that passed all spectral gates (likely unmasked shadow or haze) and vanished by the next pass. Remedied structurally with persistence confirmation (Section 4.5).
- Owner keyword matching required word-start boundaries: the surname `STREETT` substring-matched `TREE`. Prefix matching is retained so `EXCAVAT` still matches `EXCAVATING`.
- Spring tillage was the dominant residual false positive: hundreds of flags per scene date in late April, clustered at `dNBR 0.27–0.35` across working farms. Freshly disked soil mimics a light burn in `dNBR` but not in absolute `NBR`: tilled bare soil rests near `+0.05` to `+0.15`, while char sits at or below zero. Raising `DNBR_MIN` to `0.32` and lowering `POST_NBR_MAX` to `0.0` removed the tillage band on physical grounds.

Expected calibrated behavior: dozens of legitimate small legal burns will still flag during peak spring burn season at scores 1–2 and sink to the bottom of the queue; multi-date recurring sites and keyword-matched owners rise to the top. Summer detections are disproportionately valuable — legal agricultural burning largely ceases once crops are planted, so mid-year burn signatures skew toward disposal operations.

## 6. Adapting the System to a New County

The pipeline contains no Clinton-County-specific logic. Adaptation is a configuration exercise:

4. **Bounding box.** Set `COUNTY_BBOX` to `[min_lon, min_lat, max_lon, max_lat]` in WGS84, padded slightly beyond the county line. Derivable from the parcel layer's extent.
5. **Tile grid.** Each Process API request must stay under 2,500 px per side at 10 m resolution ( $\approx 25$  km). Set `TILE_GRID` so `bbox_side / TILE_GRID  $\leq$  ~24` km. A typical Midwest county fits a 2×2 grid; large western counties may need 3×3 or 4×4.
6. **Parcel layer.** Obtain parcels as WGS84 GeoJSON and map the three field names in `config.py` (Section 3.2).
7. **Keyword lists.** Tune `OWNER_KEYWORDS / OWNER_HOT_KEYWORDS` to the trade being monitored — tree waste, demolition debris, tire disposal, etc.
8. **Credentials.** Each deployment should use its own free Copernicus account and OAuth client.
9. **Calibrate.** Run the Section 5 procedure. Land-use mix (row crop vs. pasture vs. timber), soil color, and burn culture differ by county; expect one or two threshold iterations.

Regulatory framing also transfers: in Missouri, open burning of trade waste is restricted under 10 CSR 10-6.045, and analogous rules exist in most states. Verify the local rule before characterizing any detection as a potential violation.

## 7. Installation and First Run

10. Install Python 3.11+ and dependencies: `pip install numpy scipy shapely requests tiffio`

11. Register at [dataspace.copernicus.eu](https://dataspace.copernicus.eu) → verify email → open [shapps.dataspace.copernicus.eu/dashboard](https://shapps.dataspace.copernicus.eu/dashboard) → User settings → OAuth clients → Create. Copy the client ID and the one-time secret.
12. Enter credentials in `config.py` (or set `CDSE_CLIENT_ID` / `CDSE_CLIENT_SECRET` environment variables, which take precedence — preferred for any shared or cloud deployment).
13. Place `parcels.geojson` beside `run.py`.
14. Verify: `python run.py --test-auth` → “CDSE auth OK”
15. Backfill: `python run.py --backfill YYYY-MM-DD` (3 months recommended). The first clear scene per tile seeds the baseline; detection begins with the second.
16. Open `review_map.html` and run the calibration loop (Section 5).

## 8. Automated Weekly Operation

**GitHub Actions (recommended).** Push the folder to a private repository — private is mandatory, because detections carry unverified owner names and the repo may carry credentials. Add `CDSE_CLIENT_ID` and `CDSE_CLIENT_SECRET` as Actions secrets. The included workflow (`.github/workflows/burnwatch.yml`) runs every Monday morning, installs dependencies, processes all scenes since the last run, and commits the updated database, `detections.geojson`, and `review_map.html` back to the repository. It can also be triggered manually from the Actions tab. Cost: ~2–3 minutes/week of the free Actions allowance; the host PC never needs to be on.

**Windows Task Scheduler (local alternative).** Schedule `run_weekly.bat` weekly; it changes to its own directory, runs the scan, and appends output to `state\weekly_log.txt`. The machine must be on at the scheduled time.

Both modes are self-healing: `run.py` resumes from the last processed scene date, so missed weeks, interruptions, and cloudy passes are recovered automatically on the next run. Every completed tile is committed to SQLite as it finishes.

## 9. Outputs and Data Model

Output	Format	Use
<code>review_map.html</code>	Self-contained HTML (Leaflet)	Weekly triage: ranked queue, satellite basemap, per-detection popups. Opens from disk; regenerated every run.
<code>detections.geojson</code>	GeoJSON polygons, WGS84	GIS interchange: loads directly into QGIS, Leaflet/Flask apps, ArcGIS, Google Earth. Properties: date, area, dNBR, parcel, owner, recurrence, score.
<code>state/burnwatch.sqlite</code>	SQLite	Database of record. Tables: scenes (processing ledger — proof of coverage dates), sites (recurrence registry), detections (full attributes + geometry).
<code>state/baseline_*.npy</code>	NumPy arrays	Rolling per-tile NBR baselines (internal).

python run.py --export-only regenerates the GeoJSON and map from the database at any time without fetching imagery.

## 10. Established-Pit Detection: the Three-Stage Architecture

Change detection has a structural blind spot: it sees ground BECOMING char, so a pit already established before monitoring begins produces no signal. The primary mission — finding existing pits — therefore uses a dedicated three-stage pipeline layered on the same data:

17. **Stage 1 — multi-signal sweep (scar\_sweep.py)**. Composites the most recent clear scenes and flags pit-scale ground (0.1–4 ac) that NEVER GREENS (max NDVI  $\leq$  0.35 while the county peaks at 0.6–0.8) and stays spectrally non-vegetation (MEAN NBR  $\leq$  0.18 — the mean, not the max, because a working surface cycles through char, staged wood, and ash). Elongated trench shapes are explicitly allowed and tagged; temporal variance of the surface (a working pit changes, a rooftop does not) drives ranking. Reference county yield: ~3,500 candidates.
18. **Stage 2 — evidence fusion (fuse.py)**. Each candidate is credited for independent corroboration within 250 m: confirmed burn events from change detection (+2.0 per distinct date), watchlist events (+1.5), provisional detections (+0.5). Persistent scars that also burn repeatedly rise to the top; in the reference county, 1,256 of 3,505 candidates carried convergent evidence.
19. **Stage 3 — vision classification (vision\_review.py)**. The top fused candidates are classified on Missouri statewide 6-inch aerial imagery (MSDIS ImageServer) by the Claude vision API, which judges the WHOLE visible feature — scar, ash, staged debris, haul roads, equipment — rather than pixel spectra. Output is a visual review page; a human reviewer confirms or rejects each machine verdict (reject.py records the decision and reason). Cost for a 251-candidate county run: roughly \$2–4 of API usage; everything else in the system is free.

Division of labor: the sweep finds every persistent candidate regardless of when it was established; fusion orders them by convergent evidence; vision separates pits from the dark-ground look-alikes (rooftops, feedlots, worked soil) that spectral gates cannot; the human review is final. Meanwhile the weekly change-detection net catches NEW pits as they appear, and the watchlist builds dated activity records on every confirmed site.

## 11. Validation Against Ground Truth

The complete system was validated against a known, independently documented open burn pit: a boom-fed linear trench operation of roughly 1–3 acres with a purpose-graded perimeter haul road, on a 10-acre agricultural-with-homestead parcel in Clinton County, Missouri.



*Ground-truth reference site: crescent-shaped boom-fed disposal trench with graded perimeter haul road, mixed char/ash/staged-debris surface. Aerial imagery via Google Earth, accessed July 2026.*

#### **Validation results, in sequence:**

- Watchlist time series (dark-core NBR over the site, every scene since April 5, 2026): the site was already at char level at monitoring start (49–56% of the AOI at  $\text{NBR} \leq 0$  in early April — confirming the change-detection blind spot for pre-existing pits); a major burn event registered the week of May 12 (core NBR  $-0.325$ , minimum pixel  $-0.704$  — unambiguous fresh char); additional probable events mid-April and mid-June; and a persistently dark pit floor visible in every subsequent clear scene.
- Blind sweep detection: the county-wide multi-signal sweep, given no coordinates, produced a candidate 62 m from the site, on the correct parcel.
- Blind vision classification: the 6-inch aerial chip of the site was classified PIT at 82% confidence, with the model independently reporting: crescent-shaped disturbed area, ash deposits, char, staged woody debris piles, haul road — matching the ground-truth feature set item for item.

**County census outcome (reference run, July 2026):** 420 mi<sup>2</sup> → 3,505 spectral candidates → 251 vision-classified → 8 machine-flagged → 4 human-verified burn pits, including the ground-truth site (found blind) and one discovery carrying five confirmed 2026 burn dates. Machine false positives were dominated by one class — active pond/lagoon construction, whose excavated dark earth, machinery, and converging tracks mimic pit features — all removed in minutes of human review; recommended prompt hardening is noted in `vision_review.py`.

## 12. Limitations and Responsible Use

- Temporal resolution. ~5-day revisit minus cloud cover yields an effective clear look every 1–2 weeks. Scars persist spectrally for weeks, so established burn sites are caught reliably; a single small burn fully extinguished and cleaned between passes can be missed. NASA FIRMS thermal alerts (free API) are a complementary same-day feed worth adding for active-fire notification.
- Spatial resolution. At 10 m, the minimum reliable detection is roughly 0.1 acre of contiguous burn signature. Small single-pile burns fall below the floor by design.
- Attribution. Detections locate events on parcels; they do not identify actors or materials. Parcel ownership is the deed holder, who may not be the burner.
- False positives persist at reduced rates: legal open burning, harvest/land-clearing edge cases, and unusual soil exposures. The score ranking exists precisely because verification capacity is finite.

**Publication discipline.** Never publish raw detections with owner names attached; an unverified flag naming an individual is defamation exposure. The defensible workflow is: flag → ground verification (dated photographs from public rights-of-way) → regulatory referral citing coordinates and the specific Sentinel-2 acquisition dates → publication, if warranted, only of verified findings. In referrals, cite each detection as a dated ESA Sentinel-2 L2A observation; scene dates bound the event window (“at or before this pass, since the previous clear pass”) rather than pinpointing the burn moment.

## Appendix A. Calibrated Configuration Reference (Clinton County, MO)

```
COUNTY_BBOX      = [-94.615, 39.445, -94.195, 39.755]
GSD_M            = 10          # meters/pixel
TILE_GRID        = 2          # 2x2 tiles, ~1800 px per side
MIN_CLEAR_FRACTION = 0.25
MAX_SCENE_CLOUD_PCT= 85

DNBR_MIN         = 0.32      # calibrated up from 0.27 (tillage band)
POST_NBR_MAX     = 0.0       # calibrated down from 0.10 (char, not bare soil)
MIN_BURN_ACRES   = 0.08
MAX_BURN_ACRES   = 5.0
MERGE_RADIUS_PX  = 3
MIN_COMPACTNESS  = 0.12
POST_NIR_MAX     = 0.18      # char is NIR-dark; cut hay/stubble is bright
CONFIRM_NBR_MAX  = 0.15      # scar persistence threshold on second look
CONFIRM_MIN_CLEAR = 0.5      # min clear fraction to attempt resolution
SITE_MATCH_RADIUS_M= 250

PIT_NDVI_MAX     = 0.35      # never-greens gate (sweep)
PIT_NBR_MEAN_MAX = 0.18      # persistent non-vegetation gate (sweep)
FUSE_W_CONFIRMED = 2.0       # fusion credit per confirmed burn date
VISION_TOP_N     = 250       # candidates sent to vision stage
SMALL_PARCEL_ACRES = 40.0
```

Weekly cadence, county-wide coverage, and zero marginal data cost make this class of monitoring reproducible by any civic group with one technical volunteer. Total recurring cost of the reference deployment: **\$0**.