solar_panel_telemetry.py — Full Documentation

1) Purpose & Scope

This script generates synthetic, per-panel telemetry for a solar PV site. It produces realistic time-series data with a diurnal irradiance curve, cloud variability, panel-to-panel differences, temperature effects, slow degradation, and occasional faults. Output can be CSV (default) or JSONL for easy ingestion into streaming pipelines, databases, or dashboards.

---

2) Quick Start

2.1 Install & Run

No external dependencies (Python 3.8+).

python solar_panel_telemetry.py --panels 100 --hours 12 --step 60 --out telemetry.csv

2.2 Common Examples

- 10 panels, 1 hour, sample every 30 s to stdout (CSV):

python solar_panel_telemetry.py --panels 10 --hours 1 --step 30

- JSONL output for 5 panels, 30 minutes, every 10 s:

python solar_panel_telemetry.py --format jsonl --panels 5 --minutes 30 --step 10 --out data.jsonl

- Fixed start time and daylight window:

python solar_panel_telemetry.py --start "2025-10-18T06:00:00" --hours 8 --daylight 05:45-18:30 --seed 123

---

3) Command-Line Interface (CLI)

Flag	Type	Default	Description
--panels	int	50	Number of panels to simulate.
--start	str (ISO 8601)	now (UTC)	Start timestamp (e.g., 2025-10-18T06:00:00). If timezone is included, it's converted to UTC.

Flag	Type	Default	Description
--hours	float	–	Duration in hours (mutually exclusive with --minutes). If neither --hours nor --minutes specified, it defaults to 8 hours.
--minutes	float	–	Duration in minutes (mutually exclusive with --hours).
--step	int (seconds)	60	Sampling period between points.
--out	str	-	Output path or - for stdout.
--format	csv | jsonl	csv	Output file format.
--seed	int	42	RNG seed for reproducibility.
--site	str	Site-A	Logical site name (label only; not emitted in records).
--daylight	HH:MM-HH:MM	06:00-18:00	Daylight window driving irradiance envelope.
--timezone	str	UTC	Label only; timestamps are emitted in UTC (ISO with Z).

Notes

- Provide **only one** of --hours or --minutes. If both are omitted, duration is 8 hours.
- Timestamps are ISO 8601 UTC (e.g., 2025-10-18T06:01:00Z).

---

4) Output Schema

4.1 CSV Columns (default) & JSONL Keys

Field	Type	Units	Description
timestamp_utc	string	ISO8601	UTC timestamp (...Z).
panel_id	string	–	Unique panel identifier (P00001, ...).
string_id	string	–	Electrical string grouping (e.g., S01).

Field	Type	Units	Description
status	enum	–	OK, WARNING, or FAULT.
fault	enum	–	NONE, SHADING, HOTSPOT, STRING_OPEN, INVERTER_TRIP, SOILING.
power_w	float	W	DC power at panel output (capped ~110% of STC).
voltage_v	float	V	Approx. MPPT-region voltage with noise.
current_a	float	A	Derived from power_w / voltage_v.
irradiance_wm2	float	W/m^2	Site/string-level irradiance proxy (0 at night).
ambient_temp_c	float	°C	Ambient air temperature (diurnal model).
cell_temp_c	float	°C	Estimated cell temperature (NOCT-style).
orientation_deg	float	°	Panel azimuth (≈ south = 180°).
tilt_deg	float	°	Panel tilt.

4.2 Sample Records

CSV

timestamp_utc,panel_id,string_id,status,fault,power_w,voltage_v,current_a,irradiance_wm2,ambient_temp_c,cell_temp_c,orientation_deg,tilt_deg

2025-10-18T06:00:00Z,P00001,S01,OK,NONE,12.44,35.12,0.354,46.8,21.73,22.17,178.4,24.8

JSONL

{"timestamp_utc":"2025-10-18T06:00:00Z","panel_id":"P00001","string_id":"S01","status":"OK","fault":"NONE","power_w":12.44,"voltage_v":35.12,"current_a":0.354,"irradiance_wm2":46.8,"ambient_temp_c":21.73,"cell_temp_c":22.17,"orientation_deg":178.4,"tilt_deg":24.8}

---

5) Physical & Statistical Modeling

5.1 Diurnal Irradiance

- Daylight window parsed from --daylight (e.g., 06:00-18:00).

- Envelope uses a **smooth bell** (sin(pi*x)^{1.5}) with **quintic smoothstep** edges to mimic sunrise/sunset.
- Outside daylight: irradiance = 0 → power ≈ 0.

5.2 Clouds & Local Variability

- Per-timestamp **site cloud factor** (0.6–1.0) mixes frequent light attenuation with occasional dips.
- Per-string jitter around site cloud factor to emulate partial cloud cover.

5.3 Temperature Model

- Ambient temperature varies ±8 °C daily, peaking mid-afternoon.
- **Cell temperature** ≈ ambient + (irradiance/800)*20°C (simple NOCT-like rule).

5.4 Panel Characteristics & Diversity

Each panel has:

- p_stc_w (≈370–430 W), v_mppt (≈33–40 V), current scale, temperature coefficient (−0.45% to −0.35% per °C).
- **Slow degradation** per day (≈0.003–0.007%/day).
- **Orientation/tilt modifier** (0.85–1.05×) favoring south and moderate tilt.
- Small **efficiency jitter** (±2%).

5.5 Faults & Effects

- Random fault assignment with low probabilities; most time NONE/OK.
- Effects:
  - SHADING: reduce power 10–40%.
  - HOTSPOT: reduce power 20–60% and bump cell_temp_c by +3..+8°C.
  - STRING_OPEN / INVERTER_TRIP: force power_w = 0, current_a = 0.
  - SOILING: reduce power 5–20%.

---

6) Code Architecture

6.1 Key Data Structures

- **PanelSpec** (@dataclass): per-panel immutable spec (STC power, MPPT voltage, temp coeff, etc.).

6.2 Core Functions (Public Behavior)

- **generate_panel_fleet(n: int, seed: int) -> List[PanelSpec]**
  Creates n panels with randomized but realistic specs (deterministic via seed).
- **make_time_range(start: datetime, duration: timedelta, step_s: int) -> Iterator[datetime]**
  Yields timestamps from start to start+duration inclusive at step_s intervals.
- **compute_telemetry(spec, ts, dl_start, dl_end, day_index, cloud_factor) -> Dict[str, object]**
  One record for one panel at time ts. Combines irradiance, temperature, orientation, degradation, noise, and faults to produce the output schema.

6.3 Supporting Utilities

- **Daylight / Diurnal**: parse_daylight_window, diurnal_irradiance_factor, smoothstep.
- **Environment**: ambient_temperature_c, panel_cell_temp_c, cloud_cover_factor.
- **Panel Modifiers**: orientation_tilt_modifier.
- **Faults**: assign_fault.
- **Time Helpers**: utc_now_truncated, seconds_since_midnight, day_number_since.

6.4 Main Flow

1. Parse CLI args; set seed.
2. Determine start_utc, duration, daylight window.
3. Create fleet of PanelSpec.
4. For each timestamp:
   - Draw site cloud factor; derive per-string factor.
   - For each panel: call compute_telemetry(...).
   - Emit record (CSV row or JSON line).

---

7) Determinism & Reproducibility

- **--seed** controls RNG across fleet creation, clouds, noise, and faults.
- Given the same --seed, arguments, and Python version, you'll get identical outputs.

---

8) Performance Considerations

8.1 Throughput

- Pure-Python; no heavy deps. On a modern laptop, generating **millions** of rows is reasonable.
- Tips for high volume:
  - Use **CSV** (faster than JSONL).
  - Increase --step (fewer rows).
  - Write to a **local SSD** path (avoid stdout for huge runs).
  - Consider sharding runs by time window and panel groups.

8.2 Memory

- Streamed generation (records are not buffered). Memory footprint remains low.

---

9) Extending the Model

9.1 Add New Fault Types

1. Edit FAULT_CATALOG to add (name, status, probability) tuples.
2. Extend the if fault != "NONE": block in compute_telemetry with the effect.

9.2 Per-Site Configuration

- Inject per-site constants (e.g., daylight/seasonality) via CLI or a small config loader.
- Add --latitude/--longitude and compute daylight from date if needed.

9.3 AC Side / Inverter Modeling

- Add DC→AC curve and clipping at inverter rating.

- Emit additional metrics (frequency, power factor, inverter state).

9.4 Weather Inputs

- Replace synthetic ambient_temperature_c and clouds with a real feed or a weather file.

9.5 Multi-File Partitioning

- Add --partition hourly|daily|string to rotate files by time or string ID.

---

10) Function-by-Function Reference

generate_panel_fleet(n, seed) -> List[PanelSpec]
Creates n panels. Each has randomized:

- Power at STC (p_stc_w), MPPT voltage (v_mppt), current scale, temp coeff, daily degradation, efficiency jitter, azimuth/tilt, and a string_id (20 panels per string by default).

diurnal_irradiance_factor(ts, dl_start, dl_end) -> float

- Returns 0..1 daylight factor with smooth sunrise/sunset edges and midday peak.

cloud_cover_factor() -> float

- Random multiplicative attenuation (≈0.6–1.0).

ambient_temperature_c(ts, base=26.0) -> float

- Daily sinusoid around base with amplitude 8 °C (peak mid-afternoon).

panel_cell_temp_c(ambient_c, irradiance_wm2) -> float

- NOCT-style approximation: ambient + (irr/800)*20.

orientation_tilt_modifier(orientation_deg, tilt_deg) -> float

- Favors southish (≈180°) and tilt ~25°. Produces 0.85–1.05× power modifier.

assign_fault() -> (fault: str, status: str)

- Draws from FAULT_CATALOG. Low overall probability; otherwise returns ("NONE","OK").

compute_telemetry(spec, ts, dl_start, dl_end, day_index, cloud_factor) -> dict

- Produces the **single output record** combining all models and clipping/rounding.

---

11) Assumptions & Limitations

- **Simplified physics**: No spectral effects, shading geometry, module IV-curve, or wind cooling.
- **Irradiance** is an abstract factor not derived from sun position/clear-sky models.
- **Fault probabilities** are illustrative, not based on field failure rates.
- **All timestamps** emitted in UTC; the --timezone flag is a label only.

---

12) Troubleshooting

Symptom	Likely Cause	Fix
"Provide ONLY one of --hours or --minutes."	Both given simultaneously.	Supply only one, or neither for default 8 hours.
All zeros throughout the day	Daylight window doesn't cover --start range	Adjust --daylight, --start, or extend duration.
Very small currents at night but non-zero voltage	Expected (electronics idle noise).	Working as designed.
Non-deterministic outputs	Missing/changed --seed.	Set --seed and keep environment constant.
