

# LCA Tool — Output CSV Examples & Field Explanations (No-Emulator Mode)

This report contains concrete example rows for each `record_type` in the output CSV `scenario_output.csv` and a detailed, field-by-field explanation for each row. These examples assume NO emulator is used — final metrics are produced by a deterministic LCA engine (Brightway2/OpenLCA) with optional Monte Carlo sampling. Imputed inputs are documented in PROVENANCE rows. Use these examples as copy/paste-ready templates for SIH and documentation.

## 1. SCENARIO\_META (one row)

CSV example row:

```
SCENARIO_META,scen-AL-001,"Al Sheet Baseline","1 kg Al sheet","kg","cradle-to-cradle","IN",2025,"IPCC2013_GWP100a",1000,"recycled_content",false,"run-20250916-001",done,"2025-09-16T12:00:00Z","2025-09-16T12:00:00Z",false,""
```

Field-by-field explanation:

**record\_type:** 'SCENARIO\_META' — fixed token identifying this row as scenario metadata.

**scenario\_id:** scen-AL-001 — unique scenario identifier.

**scenario\_name:** 'Al Sheet Baseline' — human readable title for the scenario.

**functional\_unit:** '1 kg Al sheet' — the reference unit for all results.

**functional\_unit\_base:** kg — canonical unit for normalization.

**system\_boundary:** cradle-to-cradle — scope; chosen by user (not imputed).

**geography:** IN — regionalization used for grid and transport factors.

**year:** 2025 — study reference year.

**lcia\_method:** IPCC2013\_GWP100a — LCIA method used (user choice; not imputed).

**allocation\_method:** recycled\_content — allocation rule chosen by user (not imputed).

**currency / units\_system / author:** metadata fields.

**run\_id:** run-20250916-001 — backend job id (generated).

**run\_status:** done — run completed successfully (final deterministic metrics present).

**mc\_samples:** 1000 — Monte Carlo sample count used to propagate uncertainty.

**run\_started\_at / run\_finished\_at:** ISO timestamps for run timing.

**emulator\_used:** false — explicit: NO emulator used; deterministic engine produced final metrics.

**notes\_results:** free-text summary; mention imputed inputs & provenance IDs (e.g. prov-002).

## 2. PROCESS (one row)

CSV example row:

```
PROCESS,scen-AL-001,p_smelt_01,"Primary Smelting - Plant A",smelting,"Hall-Hérault","IN-OR",1,15000
```

Field-by-field explanation:

### Inputs (user-provided or accepted):

- **process\_id**: p\_smelt\_01 — unique identifier for this process.
- **process\_name**: 'Primary Smelting - Plant A' — human label.
- **process\_type**: smelting — grouping for templates and ML.
- **technology**: 'Hall-Hérault' — recommended; can be ML-predicted if missing (lower confidence).
- **location**: IN-OR — regionalization for grid and transport factors.
- **functional\_unit\_multiplier**: 1 — how many FU this process yields per operation (required).
- **plant\_capacity\_tpa**, **plant\_efficiency**, **labour\_hours\_per\_FU**, **labour\_cost\_per\_hour**, **capex\_annuity\_per\_FU**, **opex\_per\_FU**: operational & economic metadata (some imputation-eligible).
- **data\_source**, **confidence\_score**, **notes**: provenance & Q information.

### Appended computed fields (deterministic LCA results):

- **GWP\_contribution\_kgCO2e**: 9.82 — sum of this process's flows × CFs + allocated upstream burdens. Not directly imputed (but depends on any imputed flows).
- **Energy\_contribution\_MJ**: 51.70 — total energy attributed to process per FU (derived).
- **Water\_contribution\_m3**: 0.32 — water consumption for this process per FU (derived).
- **Circularity\_contribution**: 0.12 kg/FU — mass recovered; imputer may predict recycling yields if missing.
- **process\_uncertainty\_ci\_low / ci\_high**: 8.00 / 11.60 — 95% CI for process GWP from MC traces.
- **notes\_result**: explanatory caveats (e.g., electricity was imputed; check meters).

### 3. FLOW — Electricity (imputed)

CSV example row:

```
FLOW,scen-AL-001,f_elec_01,p_smelt_01,"electricity - grid",input,14.36,kWh,energy,,,,,"predicted",0
```

Field-by-field explanation:

#### Core fields:

- **flow\_id:** f\_elec\_01 — unique flow id.
- **process\_id:** p\_smelt\_01 — FK to the PROCESS row.
- **flow\_name:** 'electricity - grid' — mapping to energy flow.
- **direction:** input — consumed by the process.
- **quantity:** 14.36 kWh per FU — THIS WAS IMPUTED by the imputer (measurement\_type='predicted').
- **unit:** kWh (backend will normalize to MJ for energy calculations).
- **flow\_category:** energy.

#### Imputation details:

- **imputation\_method:** XGBoost\_imputer\_v1.5 — model used to predict the quantity.
- **confidence\_score:** 0.62 — model confidence.
- **notes:** 'Predicted by imputer; accept/edit/reject in UI' — UI must prompt accept/edit/reject.

#### Appended computed fields:

- **GWP\_contribution\_kgCO2e:** 8.62 — computed as quantity × CF for regional electricity.
- **Energy\_contribution\_MJ:** 51.70 — 14.36 kWh × 3.6.
- **ci\_95\_low / ci\_95\_high:** 7.85 / 9.45 — 95% CI for the flow's GWP contribution (propagated from imputer CI).
- **provenance\_id:** prov-002 — link to the PROVENANCE row documenting the imputation.

## 4. FLOW — Alumina (reference dataset)

CSV example row:

```
FLOW,scen-AL-001,f_alumina_01,p_smelt_01,"alumina",input,1.90,kg,material,, "Al2O3",road,120,0.00,, ,
```

Field-by-field explanation:

### Core fields:

- **flow\_id:** f\_alumina\_01.
- **process\_id:** p\_smelt\_01.
- **flow\_name:** 'alumina'.
- **direction:** input.
- **quantity:** 1.90 kg per FU — provided by reference dataset (not imputed).
- **unit:** kg.
- **flow\_category:** material.

### Context / provenance:

- **data\_source:** ecoinvent:Alumina\_v3 — this flow pulls upstream embodied impacts from a reference dataset.
- **transport\_mode/distance:** road / 120 km — used to calculate transport emissions.

### Appended computed fields:

- **GWP\_contribution\_kgCO2e:** 1.20 — includes alumina production + transport burdens.
- **Energy\_contribution\_MJ:** 20.00.
- **Water\_contribution\_m3:** 0.05.
- **provenance\_id:** prov-003 — link to reference dataset provenance.

## 5. METRIC — GWP (scenario-level)

CSV example row:

```
METRIC,scen-AL-001,m_gwp,"Global Warming Potential","GWP","kg CO2e / FU",9.82,9.95,8.10,11.20,8.00,
```

Field-by-field explanation:

**Purpose:** scenario-level aggregated LCIA result for GWP.

**Fields:**

- **metric\_id:** m\_gwp.
- **metric\_name:** 'Global Warming Potential'.
- **unit:** 'kg CO2e / FU'.
- **median:** 9.82 — median of MC samples (deterministic LCA engine + MC).
- **mean:** 9.95 — arithmetic mean (optional).
- **p10 / p90:** 8.10 / 11.20 — percentiles.
- **ci\_95\_low / ci\_95\_high:** 8.00 / 11.60 — 95% CI.
- **mc\_sample\_count:** 1000 — draws used.
- **lcia\_method:** IPCC2013\_GWP100a.
- **provenance\_id:** prov-manifest-001 — points to the manifest of provenance entries for the scenario.

## 6. MC\_TRACE (sample rows)

CSV example row:

```
MC_TRACE,scen-AL-001,1,m_gwp,9.10,57431
MC_TRACE,scen-AL-001,2,m_gwp,10.05,57431
```

Field-by-field explanation:

**Purpose:** raw Monte Carlo sample outputs for reproducibility or in-depth analysis.

Each MC\_TRACE row records one draw's metric value. For large MC sets we store traces zipped and reference via `traces_url` in SCENARIO\_META.

Fields: `record_type`, `scenario_id`, `sample_id`, `metric_id`, `sample_value`, `seed`.

## 7. SENSITIVITY (OAT example)

CSV example row:

```
SENSITIVITY,scen-AL-001,smelt_elec_kWh_per_FU,14.36,11.49,17.23,8.53,11.10,OAT,"±20% OAT on electricity"
```

Field-by-field explanation:

**Purpose:** document single-parameter perturbation results (One-At-a-Time) or variance-based indices (Sobol).

**Fields:**

- **param\_name:** the parameter tested (smelt\_elec\_kWh\_per\_FU).
- **param\_baseline:** 14.36 — baseline value.
- **param\_low / param\_high:** tested low/high values (here  $\pm 20\%$ ).
- **metric\_impact\_low / metric\_impact\_high:** resulting metric values (GWP) when parameter is set to low/high and deterministic LCA is re-run.
- **method:** OAT (here).
- **note:** explains deterministic re-run and provenance.

## 8. RECOMMENDATION (deterministic verification)

CSV example row:

```
RECOMMENDATION,scen-AL-001,rec-001,"Increase recycled feed to 40% in smelter","Replace 40% of virgi
```

Field-by-field explanation:

**Purpose:** actionable, ranked recommendations with estimated benefit and uncertainty.

**Fields:**

- **rec\_id:** unique recommendation id.
- **title / description:** what to change operationally.
- **delta\_metric\_id:** which metric is impacted (m\_gwp).
- **delta\_median / delta\_ci\_low / delta\_ci\_high:** expected change and uncertainty after deterministic verification (snapshot + full LCA).
- **effort** and **confidence:** qualitative metadata.
- **how\_computed:** 'deterministic' — means recommendation was tested by running the full deterministic LCA on a scenario snapshot.



## 9. PROVENANCE (imputed input audit)

CSV example row:

```
PROVENANCE,prov-002,scen-AL-001,flows[f_elec_01].quantity,14.36,kWh,predicted,"XGBoost_imputer_v1.5"
```

Field-by-field explanation:

**Purpose:** audit trail entry for an imputed value. ISO-style traceability must show: which field was imputed, value used, unit, model, CI, SHAP explanation, who accepted it and when.

**Fields explained:**

- **provenance\_id:** prov-002 — unique id referenced by FLOW rows (provenance\_id field).
- **field\_path:** flows[f\_elec\_01].quantity — identifies the exact field imputed.
- **value / unit:** 14.36 kWh — value accepted and used in LCA.
- **source\_type:** predicted — indicates ML imputation.
- **source / imputation\_method:** XGBoost\_imputer\_v1.5 (model/version).
- **ci\_95\_low / ci\_95\_high:** 13.15 / 15.25 — imputer uncertainty bounds.
- **shap\_top:** JSON with top SHAP drivers (signed contributions).
- **last\_updated\_by / last\_updated\_ts:** who accepted and when.

## 10. Final checklist and remarks

- All imputed inputs must have PROVENANCE rows with model/version, CI, and SHAP explainability.
- Final METRIC rows are authoritative deterministic-LCA outputs (no emulator). Show median & 95% CI from MC.
- SCENARIO\_META.emulator\_used must be false (explicit) if you do not use emulator.
- UI must present imputer predictions with Accept/Edit/Reject and display SHAP reasons.
- Large MC\_TRACE sets should be stored zipped and referenced by traces\_url in SCENARIO\_META.
- Recommendations that claim final deltas should be verified by deterministic snapshot runs (how\_computed=deterministic).

Report generated for SIH LCA tool documentation. The examples show exact CSV rows and field explanations suitable for judges and developers. All values above are illustrative; replace with real data for production runs.