HIGG ASSESSMENT MODEL

GHGP Protocol Compliant FEM Calculations

Version next · 11 September 2026

This document provides guidance on the Greenhouse Gas Protocol (GHGP) compliant FEM calculations. The calculations are used to generate the data available in the Assessment, FacilityType, ReportedEnergyUse, EnergyUse, and Emissions records.

Contents

This document provides guidance on the Greenhouse Gas Protocol (GHGP) compliant FEM calculations. The calculations are used to generate the data available in the `Assessment`, `FacilityType`, `ReportedEnergyUse`, `EnergyUse`, and `Emissions` records.

The methodology for these calculations has evolved over time. This document outlines the methodology for each Higg Facility Environmental Module (FEM) cadence, starting with FEM 2017.

Purpose of the Calculations

The primary purpose of these calculations is to provide a consistent and transparent methodology for calculating greenhouse gas (GHG) emissions from facilities, in line with global best practices. The two main goals of the methodology are:

1. **GHGP Compliance:** The methodology is designed to be compliant with the Greenhouse Gas Protocol (GHGP), which is the most widely used international accounting tool for government and business leaders to understand, quantify, and manage greenhouse gas emissions.
    - **Example:** A key aspect of GHGP compliance is the dual reporting of Scope 2 emissions. The calculations provide both market-based and location-based emissions, allowing facilities to report their emissions in the manner that best represents their energy procurement choices.
2. **Methodological Consistency:** The methodology is designed to be as consistent as possible across all FEM cadences. This allows for meaningful year-over-year comparisons of a facility's environmental performance.
    - **Example:** To ensure comparability, the emission factors for specific fuel sources (e.g., coal, natural gas) are held constant across all FEM cadences. This means that any change in emissions from these sources is a direct result of changes in the facility's operations, not a change in the underlying emission factors.

Record Types

The calculations produce several different record types, each with a specific purpose.

- **`Assessment`**: This record contains the aggregated energy use and emissions data for the entire facility. It provides a high-level overview of the facility's environmental performance.
- **`FacilityType`**: This record provides a breakdown of energy use and emissions by facility type (e.g., `finalProductAssembly`, `materialProduction`). This is useful for understanding the environmental impact of different production processes within a facility.
- **`ReportedEnergyUse`**: This record contains the raw energy use data as reported by the facility, converted to MJ. This data is not modeled or attributed and is not suitable for GHGP reporting, but can be useful for data quality analysis. Alongside the MJ `value`, the record carries the figure as the facility entered it: `reported_quantity` in `reported_unit_key` (the canonical unit option key — pre-FEM-2020 cadences stored per-kind option indices in the unit answer, and those are resolved to keys here, so the unit's dimension is comparable across cadences).
- **`EnergyUse`**: This record contains the modeled and attributed energy use data, appropriate for GHGP reporting. It includes separate records for market-based and location-based accounting for scope 2 energy use.

- **`Emissions`:** This record contains the calculated emissions data, based on the `EnergyUse` records. It also includes separate records for market-based and location-based accounting for scope 2 emissions.

Energy Source Categories

`EnergyUse` and `Emissions` records carry a `category` field for grouping rows by fuel family. It is derived from the **attributed** energy source's `categories` tags in `assessments/shared/data/energy_sources.stl`, taking the first match in this order:

`category`	Covers
`biogenic`	Biogenic fuels: biomass, biogas, biodiesel, ethanol, the biogenic fraction of waste fuels.
`renewable`	Non-biogenic renewables: wind, solar, hydro, geothermal, renewable hydrogen.
`coal`	Coal and coal-derived fuels, including purchased steam and district heating modelled as coal.
`gas`	Natural gas, LNG, LPG, CNG, propane.
`liquidFossilFuel`	Diesel, petrol and fuel oils.
`gridElectricity`	Consumption attributed to purchased grid electricity: grid purchases net of EACs, the non-PPA remainder of purchased renewables, the sold-certificate share of on-site generation, and purchased chilled water.
`other`	The remaining thermal sources: non-renewable hydrogen and the fossil fraction of fabric and municipal waste.

Two points to keep in mind when charting by category:

- The category follows the *attributed* source while `energy_form` follows the *reported* source, so purchased chilled water reads `energy_form: thermal` with `category: gridElectricity`.
- Under location-based accounting every purchased-renewables row is attributed to the grid, so the whole purchase lands in `gridElectricity`; under market-based accounting only the non-PPA remainder does.

Non-energy `Emissions` rows extend the vocabulary: refrigerant releases carry `fugitive`, and wastewater rows carry `biogenic` (CO_2) or `process` (CH_4).

`gridElectricity` was split out of `other` in September 2026 so grid electricity can be charted apart from the miscellaneous thermal sources. The change applies to every cadence, so historical rows re-categorize uniformly on the next full ClickHouse publish.

Methodology

The calculations begin with the `ReportedEnergyUse` records, which are generated by converting the energy consumption data reported by facilities into MJ. These records are then used to generate the `EnergyUse` records, which are suitable for GHGP reporting. This process involves modeling and attribution, which are described in more detail below.

Market-Based vs. Location-Based Accounting

The calculations distinguish between market-based and location-based accounting for scope 2 emissions.

- **Market-based accounting:** The emissions are calculated using the emission factors of the specific energy products purchased by the facility. This method is appropriate for facilities that have contractual instruments for their energy purchases, such as EACs.

- **Location-based accounting:** The emissions are calculated using the average emission factors of the grid where the facility is located. This method is appropriate for facilities that do not have contractual instruments for their energy purchases.

Modeling Assumptions

The calculations make a number of modeling assumptions to fill in data gaps and to ensure that the data is suitable for GHGP reporting. Some of the more significant assumptions are:

- **Thermal Efficiency:** Purchased steam is assumed to have a thermal efficiency of 80%. This is a conservative estimate based on industry averages.
- **Coefficient of Performance:** Chilled water is assumed to have a coefficient of performance of 3.0. This is a conservative estimate based on industry averages.
- **Off-site Renewables:** For location-based accounting, off-site renewable energy sources such as geothermal, hydro, and wind are attributed to purchased electricity. This is because these sources are typically not owned or controlled by the facility, and their environmental attributes are not transferable.
- **Unknown Source Purchased Renewables:** When a facility reports the purchase of renewable energy but does not specify the source (e.g., `renewunknown`), the emissions are calculated using the grid-average emission factor. This is a conservative approach that prevents facilities from receiving credit for renewable energy without providing the necessary documentation to support the claim.
- **Purchased Steam from Unknown Source:** When a facility reports the purchase of steam but does not specify the fuel source used to generate that steam, the methodology assumes that the steam was generated from coal. This is a conservative approach that ensures facilities are not credited with using a cleaner fuel source for their purchased steam unless they explicitly provide the data to support it.

Emissions Factors

- **Fuel Source Emission Factors:** The emission factors for specific fuel sources (e.g., coal, natural gas) are held constant across all FEM cadences. This ensures that the emissions from these sources are calculated consistently over time.
- **Grid Emission Factors:** The grid emission factors are updated with each FEM cadence to reflect the most current IEA data available for the given model version. This ensures that the emissions from purchased electricity are calculated using the most up-to-date grid mix data.

Data Quality

The calculations include a series of data quality checks to identify potential issues with the reported data. The results of these checks are exposed in two fields: `pass_data_quality` and `data_quality_flags`.

- **`pass_data_quality`**: This is a boolean field that is set to `true` if the assessment passes all data quality checks, and `false` otherwise.
- **`data_quality_flags`**: This is a list of strings that provide more detail on the specific data quality issues that were found.

There are two types of data quality flags. Flag strings are stable and prefixed by their level (`Assessment: …` / `Field: …`), so they can be matched programmatically.

- **Assessment-level flags** apply to the whole assessment:

- **Energy range.** Total reported energy above 10^{10} MJ is `implausibly high`; below 38,574 MJ (including no energy at all) is `implausibly low`.
- **Completeness.** Any energy source answered "not tracked" (`ghgp_completion_status: incomplete`) — the rows for that source carry no quantity, so the totals under-count.
- **Refrigerant plausibility** (FEM 2023 onward, where top-ups are collected; new in HAM 0.3). Refrigerant (fugitive) scope 1 of at least 10,000 t CO_2e that also exceeds the assessment's entire energy-related scope 1 + 2 (location-based, biogenic excluded) is flagged `Assessment: Refrigerant emissions of … kg CO2e are implausibly high against energy-related emissions of … kg CO2e`. The ceiling comes from IPCC 2006 Guidelines Vol. 3 Ch. 7 Table 7.9: the largest industrial refrigeration systems carry ~10 t of refrigerant and leak 7–25% a year, about 10,000 t CO_2e/yr at the highest GWPs — the most one very large system can release, and a facility running several of them has an energy footprint far above that. Empirically, below the ceiling refrigerant emissions run 1–10% of a facility's energy emissions; the flagged assessments run 15× to 20,000× (a kilogram entered as a tonne, or a charge entered as a top-up). The flag removes ~11 Mt of spurious scope 1 from FEM 2024 and under 1 Mt from other cadences. The fugitive rows are still published, so a consumer can inspect them.

- **Field-level flags** apply to specific `ReportedEnergyUse` and `EnergyUse` rows and roll up into the assessment's list:
  - **Negative quantity.** A reported energy quantity below zero.
  - **EAC / renewables double entry.** Purchased renewables reported with exactly the same value as EACs — a common data-entry confusion.
  - **LNG magnitude.** LNG above 10^9 MJ, the LNG-in-m^3-read-as-MJ confusion.

`pass_data_quality` is all-or-nothing: one flag of either level fails the whole assessment, and the verdict is stamped on every record row.

Cross-Assessment Data Quality Rules

In addition to the per-assessment checks above, a set of rules evaluates each facility's assessments **together, across cadences**, so knowledge from newer assessments can flag issues in older ones. Flags from these rules land in the same `data_quality_flags` / `field_data_quality_flags` columns with the same `Assessment:` / `Field:` prefixes, so existing consumers need no changes — but note the retroactivity: **an old assessment's flags can change because a newer assessment was posted**, not because the old assessment changed.

Each applied flag is also accompanied by a **data quality finding** — the rule's identifier, the targeted record's dimensions, the flag text, and the evidence that triggered it (typically the newer assessment's figures) — published alongside the model outputs, so "why is this flagged" never requires parsing the flag string.

Current rules:

- **`gas_volume_unit`:** Liquefied/compressed gas (LNG, LPG/propane, CNG) reported as a *gaseous* volume. These fuels are bought and burned by mass; their unit lists offered volume units through FEM 2022 and are mass/energy-only from FEM 2023 on. A facility that entered the gaseous volume over-reports by roughly the gas/liquid density ratio (~600× for LNG, ~250× for LPG). The unit change alone is not enough — most facilities converted volumes correctly — so the rule fires only when the reported energy actually collapsed across the facility's volume→mass transition: the median MJ at the last volume-unit cadence must be **at least 50×** the median MJ at the first mass/energy-unit cadence, and the mass-era median must itself be at least 1,000 MJ. That floor keeps out facilities whose gas use genuinely became very small — for example, a facility that retired a natural gas boiler but kept gas-based domestic heating, down to a single propane office heater. A large drop to a still-substantial figure is the reporting artifact this rule tracks; a drop to near-zero usage is a real transition and is not flagged. The thresholds come from the observed ratio distribution across all assessments, which separates cleanly into a ~1× bulk (correct conversions) and a 100–1,000× artifact cluster. When the rule fires, every volume-unit row of that gas family is flagged, with the computed ratio and both medians in the evidence.

- `magnitude_outlier`: A single year far off the facility's own history. For a source the facility reports across three or more distinct-valued cadences, when every other year agrees within 10× and exactly one year sits at least 100× above or below the median of the others, that year's rows are flagged. Steps of roughly 1,000× are called out specifically — they are consistent with a kWh/MWh-scale entry slip. Because assessments can carry values forward from the previous year, adjacent identical values are treated as one observation, not two, when judging the facility's history.
- `assessment_magnitude_outlier`: One year's **total** reported energy far off the facility's own history — the whole-assessment companion to `magnitude_outlier`, using the same vote (three or more distinct-valued years, the others agreeing within 10×, exactly one at least 100× off) over the assessment's total instead of a single source. It exists because a whole-assessment slip — every figure entered at the wrong scale — spreads the error across sources, so no single source may have enough history to vote on its own. The outlier year receives an assessment-level flag; roughly-1,000× steps are again called out as consistent with a kWh/MWh-scale entry slip, and value-carry-forward years are collapsed to one observation. A facility whose totals genuinely swing more than 10× between years has no self-consistent history to vote with and is not flagged. This rule deliberately overlaps the row-level rules: a year flagged both as a whole and on specific rows is the strongest signal of all.
- `site_country_changed`: The facility's site country changes between cadences. A facility is a physical site, so a changed country is either an identity problem or a data-entry error — and it matters for the numbers, because the grid emission factors follow the country. Assessments whose country differs from the facility's usual country receive an assessment-level flag.
- `steam_volume_conversion`: Purchased steam reported by volume in FEM 2017–2019, calibrated per facility. Those cadences convert steam volumes as steam vapor (6.3 MJ/m³). Many facilities' meters actually read liquid-equivalent volume (condensate or feedwater), which understates the energy by a factor of several hundred. Steam is reported by mass or energy from FEM 2020 on, and that later reporting calibrates the earlier years: where it implies liquid-equivalent metering (~1,000 kg/m³), the volume-era rows are flagged with the facility-specific understatement factor; where it is consistent with vapor volumes, no flag is raised; where no later steam reporting exists, the rows are flagged as ambiguous, since the interpretation cannot be decided from that facility alone.
- `phantom_source`: A source that appears in exactly one year of a facility's history *and* adds energy on top of the facility's usual total. A source reported once is often legitimate — a fuel switch or a new installation — and in those cases the year's total stays roughly flat because the newcomer replaces something else; those are not flagged. When the total rises by roughly the newcomer's own contribution, the energy was added rather than substituted, which suggests double-counting or misattribution. This rule deliberately overlaps with `magnitude_outlier`: two flags pulling at the same thread emphasize the highest-weight issues when many facilities are viewed together.

Cadence Specific Guidance

FEM 2017

The FEM 2017 calculations are the simplest. The energy use is calculated by converting the reported energy consumption to MJ. The emissions are then calculated by applying an emission factor to the energy use.

FEM 2020

The FEM 2020 calculations build on the FEM 2017 methodology. The main change is the addition of purchased renewable energy (`purchrenew`) to the list of off-site renewable energy sources that are attributed to purchased electricity for location-based accounting.

FEM 2021 & FEM 2022

The FEM 2021 and FEM 2022 calculations are identical to the FEM 2020 methodology. There are no significant changes in the calculation methodology for these cadences.

FEM 2023

The FEM 2023 calculations introduce a more detailed methodology for attributing energy use and emissions. This is primarily due to the introduction of more granular data from the FEM assessment.

The main changes in FEM 2023 are:

- **Steam Energy Density:** The energy density of purchased steam is calculated based on the facility-provided temperature and pressure. This allows for a more accurate calculation of the energy use from steam.
- **Energy Attribute Certificates (EACs):** The share of purchased electricity that is covered by EACs is calculated. This is used to attribute the energy use and emissions from purchased electricity more accurately.
- **Energy Source Mixes:** The calculations now account for energy source mixes, such as diesel/biodiesel blends and ethanol/petrol blends. This allows for a more accurate attribution of energy use and emissions from these sources.
- **On-site Renewable Energy:** The calculations account for the sale of Renewable Energy Certificates (RECs) from on-site renewable energy generation.

FEM 2024

The FEM 2024 calculations introduce further refinements to the methodology.

- **Biomass Certificates:** The calculations now account for the purchase of biomass certificates.
- **Generated Biomass:** The calculations now account for the on-site generation of biomass energy.

FEM 2025, FEM 2026, & FEM 2027

The FEM 2025, FEM 2026, and FEM 2027 calculations are identical to the FEM 2024 methodology. There are no significant changes in the calculation methodology for these cadences.

The `FacilityType` record gains EECI fields in FEM 2026 (and carries through to FEM 2027) — see `guidance-eeci.md` for details. This is purely additive on the FacilityType record and does not alter the underlying GHGP energy or emissions calculations documented in this file.