



CDP Full GHG Emissions Dataset

Technical Annex II: Bottom-Up Modelling Primer



Contents

1	Introduction.....	2
2	Method	3
3	Model framework.....	3
4	Data sources and granularity.....	4

1 Introduction

The Full GHG Emissions Dataset provides CDP's investor members and other stakeholders with the most up-to-date, accurate and comparable corporate GHG emissions and energy-use data. This is one of a series of documents outlining how the raw reported data is enhanced. All are available on [CDP's website](#).

- ▼ CDP Full GHG Emissions Dataset: Summary 2019
- ▼ Technical Annex I: Data Cleaning Approach
- ▼ Technical Annex II: Bottom-up Modelling Primer
- ▼ Technical Annex III: Statistical Framework
- ▼ Technical Annex IV: Scope 3 Overview and Modelling

Bottom-up (BU) modelled estimation is based on the combining of physical activity indicators (tonnes, barrels, etc.) and their associated emission factors. Because such indicators relate directly to the emitting activity, the bottom-up estimates are expected to have a narrower margin of error than the statistical estimates. Activities and products are accounted for individually, so the method is limited to homogenous sectors. Homogenous sectors are so-called because they are structured around one, or a low number of, processes and products. Such structures are most common in the upstream extraction, production and conversion industries, which is why the method accounts for the most energy- and emissions-intensive sectors. Six homogenous sectors are bottom-up modelled for corporate energy and emissions estimation, and these are displayed in the table below.

Sector activity	Data input granularity	Estimated data points
Coal mining	Asset-level (mines)	S1, S2, S3 (Use of Sold Products), Fuel, SHEC*
Oil and gas extraction	Asset-level (fields)	S1, S2, S3 (Use of Sold Products), Fuel, SHEC*
Petroleum refining	Asset-level (refineries & processes)	S1, S2, S3 (Use of Sold Products), Fuel, SHEC*
Electric power generation	Asset-level (power stations & units)	S1, S2, Fuel, SHEC*
Steel production	Asset-level (steel mills & processes)	S1, S2, Fuel, SHEC*
Cement production	Asset-level (cement plants & kilns)	S1, S2, Fuel, SHEC*

*SHEC – Purchased energy in the form of steam, heat, electricity and cooling

2 Method

The generalized methodological approach is detailed in the IPCC (2006c) 'Guidelines for National Greenhouse Gas Inventories' and is described by Equation (1):

$$\text{Emissions} = \text{Activity data} \cdot \text{Emission factor} \quad (1)$$

The level of methodological complexity is represented by three tiers: tier 1 (basic), tier 2 (intermediate), and tier 3 (advanced). Tier 1 is generally designed for the application of readily available, or aggregate, company activity data with default emission factors, which are available, for example, from IPCC default parameter tables. Tier 2 and tier 3 are designed for the use of more granular activity data and emission factors and for a wider inclusion of process parameters. Tier 2 and tier 3 are referred to as higher tier methods. The analyses undertaken for BU estimation are of tier 1 or tier 2 complexity.

The same methodological principle applies to the estimation of energy demand, for which the same activity data is combined with energy intensity data. Direct and indirect energy forms are estimated and relate to scope 1 and scope 2 emissions respectively. Direct energy is measured as the company's consumption of fuel for energy purposes. Indirect forms of energy include steam, heat, electricity and cooling (SHEC), which are generated by third parties and purchased by the company for its own use.

3 Model framework

All the key heavy emitting industrial sectors are BU modelled with the use of asset-level data. The data and estimation models, therefore, are built to reflect this (see figure 1 to view the BU model's general arrangement diagram). As depicted in Figure 1, the BU model, scripted in R, is divided into a series of sequential steps starting from raw data input and ending in desired estimated output. The importance of preparing and standardizing data inputs before the analysis stage is paramount to enabling global modelling coefficients to be applied over multiple sectors in one step, while separate functions for more detailed sector specific technological assessment may be employed independently.

Furthermore, this approach compiles and standardizes, from a diverse range of third-party data sources and sector structures, a single standardized asset-level dataset covering the world's largest industrial sources of energy use and GHG emissions. In total, this unprecedented integrated dataset holds corporate, technical, geographical, material, energy, and emissions data from over 200,000 assets worldwide, owned by 25,000 corporate entities spanning up to 9 parentship levels.

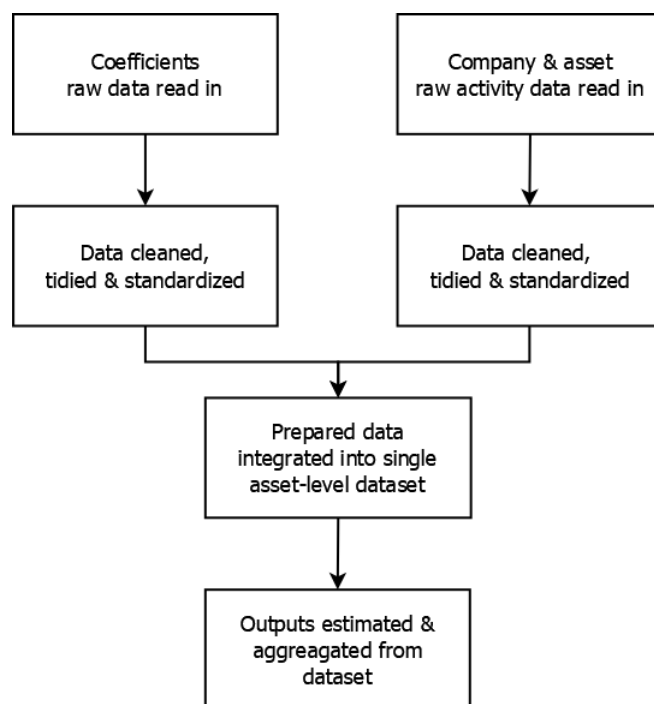


Figure 1 – General modelling process of the Bottom-up approach

4 Data sources and granularity

All asset-level data is purchased from third party data providers. Coal, oil & gas, refining, and power rely data from GlobalData, steel from Sthal's PlantFacts database, and cement from the Global Cement Directory. All sources are updated at least annually, with GlobalData undergoing continuous updates throughout the year.

Critically, the integrated asset dataset is combined with company reported environmental and activity data. This data is collected directly from annual, operational, sustainability, and integrated reports published on company websites. Further data is retrieved from US SEC 10-K and 20-F filings. This data is further supplemented by the data that companies report to CDP through the sector modules of our Climate Change questionnaire. The use of company reported activity data assures asset-data quality and coverage and provides certainty to estimates because of their direct link to the company's own data.

Collected by an expert, company reported data may be integrated with asset data at the highest level of granularity available. This meso-level of reported activity data distinguishes production activity over multiple products and technologies per sector. For example, thermal power generation is split into lignite, coal, oil, and natural gas sources, and intermediary products are incorporated, such as coke and pig iron from coke ovens and blast furnaces, respectively, in the steel sector. Meso-level coefficients are first derived by partially aggregating from the detailed technical information in the asset-level data. These are then assigned to actual meso-level production data, to provide the most accurate BU estimations.