Final Report of the Emission Factor Database Review

Prepared by the IPCC EFDB Management Group, supported by the TFI Technical Support Unit

March 2025

Contents

Table of Figures

Executive Summary

This Report on the IPCC Emission Factor Database (EFDB) has been prepared by the EFDB Management Group[1], supported by the IPCC TFI TSU[2].

Overview of Main Issues

EFDB objectives and purpose

Recent EFDB practice has emphasised those elements of the original EFDB objectives[3] that relate to offering a searchable database for use by national inventory compilers to select more appropriate EFs for their circumstances than the *IPCC Guidelines'* default EFs in order to improve the quality of their inventories.

Although data in Figures 7-9 show that the number of searches of the EFDB appear to be large and increasing over time, the EFDB has not fulfilled one critical aspect of its original objectives - that of improving national inventories in a cost-effective way. There is little or no evidence that national inventory compilers have used the searchable database to actually select EFs from the EFDB for implementation in their inventories.

The Management Review Group consider the future development of the EFDB could be reframed to emphasise other parts of the original EFDB objectives[4] and to draw on the framing for other IPCC processes[5].

Unlike the AR process, which is supported by the IPCC TG-Data, currently the TFI is not able to provide traceability for the source of default EFs in the IPCC Guidelines and, by implication, validate their accuracy and/or stability over time.

Therefore, there is a need that could be fulfilled by the EFDB if it were to be re-framed to become a repository or library of reported measurements emission/removal processes providing:

(i) transparency and traceability of default EFs reported in existing Methodology Reports of the IPCC TFI; as well as

---

[1] Takeshi Enoki and Mazhar Hayat (IPCC TFI Co-Chairs), Ole-Kenneth Nielsen and Zhu Songli (EFDB Editorial Board Co-Chairs) and Sandro Federici and Rob Sturgiss (TFI TSU).

[2] Baasansuren Jamsranjav and Lucy Garland.

[3] • Provide inventory compilers a wider choice of emission factors and other relevant parameters (including background information) than currently available in the IPCC Guidelines;
• Facilitate the selection of the most appropriate emission factors through comparisons of the underlying information; [IPCC TFI, Establishing a database on Greenhouse Gas Emission Factors, New Delhi, 2000]

[4] • Serve as a starting point for additional emission factor research programmes;
• Improve the quality and reliability of GHG inventories;
• Increase the transparency of GHG inventories;
• Provide a valuable input to assessment, review and future updates of the IPCC Guidelines [IPCC TFI, Establishing a database on Greenhouse Gas Emission Factors, New Delhi, 2000].

[5] The terms of reference for IPCC TG-Data is to "provide guidance… in order to provide curation, transparency, traceability and stability of data ..related to the reports of the IPCC" [Document: IPCC-XLVII/Doc. 9, Annex 1].

(ii) a basis for testing the accuracy/stability of IPCC default EFs over time, space and other characteristics

with the aim of building the confidence of users and stakeholders in the *IPCC Guidelines* (and by extension the quality, reliability and transparency of GHG inventories that use the *IPCC Guidelines*) and to serve as a starting point for future research and IPCC Guidelines development efforts.

Data collection

Originally, the EFDB developers envisaged that reliance would be placed on external voluntary data submissions to populate the database[6].

However, data in Figure 5 show that only around 3.5 voluntary submissions from outside experts (e.g., researchers, inventory compilers) have been received each year, on average.

EFDB Data Meetings were envisaged as a way to bolster these limited external efforts for data collection. The Data Meetings have been successful in increasing the number of datapoints for inclusion in the EFDB – but this process has not been so cost effective. There are around 1018 (out of 3179 data accepted for inclusion into the EFDB) data points accepted for the EFDB by the Data Meetings (2014-2021). The cost to the IPCC Trust Fund of this work since 2014 has been around 220,000CHF (which is budget used for journeys of participants to attend the Data Meetings) at an average cost of around 159CHF per accepted external data point on average[7] (Figure 11).

Continued reliance on voluntary submissions of EF data or collections through routine Data Meetings may no longer be seen as a cost-effective strategy.

Alternatively, the future development of the EFDB could rely on the TSU for data collection and supplemented with Data Meetings targeted at potential large dataset providers held on a case-by-case basis (like, for example, the budget provision for a Contingency Expert Meeting).

TSU data collection could begin to build on the existing database by targeting contributors to the *2019 Refinement* and with a focus on sources/sinks with the highest quality/regional/other characteristic variations: fugitive emissions, AFOLU, waste while also searching other public sources of data.

Quality Control

[6] 'The EFDB will be an open system that will contain information to be provided by its users.' (Establishing a database on Greenhouse Gas Emission Factors, Report of the First Expert Meeting, Paris, 2001)

[7] There are additional costs to the TSU in data collection that are not reflected in this calculation. Literature search/review to identify potential data, data sources and data providers is carried out mostly by TSU.

In initial planning for the EFDB emphasis was placed on the use of labour-intensive, ex-ante quality control of data points by an EFDB Editorial Board – comprising 36 members[8]:

Data in Figure 11 show that the cost to the IPCC Trust Fund of this form of intensive quality control since 2014 has been around 472,000CHF (which is budget used for journeys of Editorial Board members to attend the EFDB meetings) or around 168CHF per approved external data point on average.

Alternatively, the future development of the EFDB could place less emphasis on labour-intensive, ex ante controls and more emphasis on light-touch curation of data by the TSU supported by the use of tools to allow more systematic ex-poste population reviews of the data.

Structure of the EFDB Database

The current structure of the EFDB supports ready analysis of some of the data in the Energy Sector, but is less useful when it comes to other IPCC Sectors.

Work is currently underway to analyse EFDB EF and parameter data and to design appropriate stratifications of the database in order to maximise the utility of the data.

Once the specifications for the more detailed stratification of the database have been completed, resources will need to be found if the database is to be modified. Resources could potentially be diverted from the current resources expended on data collection and ex ante quality controls or allocated to IT inventory software development

Future EFDB outputs

The database of EF measurements is potentially a valuable asset that could be analysed to demonstrate the quality of the *IPCC Guideline* default EFs or, if necessary, to demonstrate the limitations or need for future updates.

Technical Bulletins could be the vehicle for the publication of the analysis of the data in the EFDB. Appendix 1 shows a glimpse of the information in the EF database as it currently stands and how it might be used to demonstrate (or challenge) the robust nature of the default EFs in the *IPCC Guidelines* in future.

[8] The objective of the EFDB Editorial Board is to ensure all emission factors and other parameters contained in the emission factors database (EFDB) … are scientifically sound according to the criteria endorsed by the Task Force Bureau (TFB) on the IPCC-NGGIP [Bratislava draft TOR]

Summary of Possible Recommendations

1. ***Reframe the purpose of the EFDB***
    a. The primary objective of the EFDB is to be re-framed as a repository or library of EF measurements.
        i. which could be analysed to support the transparency, traceability and accuracy/stability of IPCC default EFs.
    b. A secondary objective should be as a searchable database for users to obtain EFs for use in national inventories directly.

2. ***Reform the process of data discovery***
    a. The TSU should remain responsible for the role of the collection of data from the Editorial Board members and from others for inclusion into the EFDB.
        i. Data Meetings with specialist data providers could be held on a case by-case basis.
    b. Greater prominence should be given to the identification and collection of data by the Editorial Board.

3. ***Reform the process of quality control***
    a. Relax the emphasis on explicit, ex ante control of inputs into the database by an EFDB Editorial Board and place greater emphasis on ex-poste analysis and review of the EF data populations by the TSU in conjunction with strengthened TFB oversight.
    b. Two options for reform of ex-ante controls could be considered:
        i. Option A; Abolish the EFDB Editorial Board. The functions of the Editorial Board could be assumed by the EFDB Management Group while relying on strengthened TFB oversight and TSU curation and ex-poste analysis/review for quality control of the database as a whole. Meetings would be held in-person or online on an ad-hoc basis (for example, in conjunction with Data Meetings).
        ii. Option B: Streamline the EFDB Editorial Board – maintain the EFDB, but with a reduced number (7) of sectoral experts while relying on strengthened Management Group and TFB oversight and TSU curation and ex poste analysis/review for quality control of the database as a whole. Meetings would be held in person or online on an ad hoc basis (for example, in conjunction with Data Meetings).
            1. Revised EFDB Editorial Board Terms of Reference and EFDB Procedures including procedures as to how to deal with large datasets are attached.
    b. The effectiveness of the adopted option should be reviewed by the Management Group after 2 years.

4. Reform the EFDB structures

a. Implement the detailed stratification of the database currently under preparation to enhance the utility of the database; and
b. Explore the possibility of implementing an API (Application Programming Interface) to allow other software to make searches into the database and retrieve data, subject assessment of legal and process considerations as well as available financial resources.

5. Streamline the process of data upload into the EFDB

a. Implement the simplification and optimization of the steps involved in data upload including the data entry forms to make the data upload process faster and more efficient.
b. Explore the possibility of a web-based interface which may provide feedback to the compiler during the compilation of data, subject to available financial resources.

6. Reform the planned EFDB outputs

a. The TSU could prepare Technical Bulletins to communicate analysis of the EFDB data which could provide the evidence and basis for confidence in the IPCC default EFs or, if necessary, provide a basis for future updates.

b. Searchable database for external users would be maintained but with a more sophisticated stratification of available data.

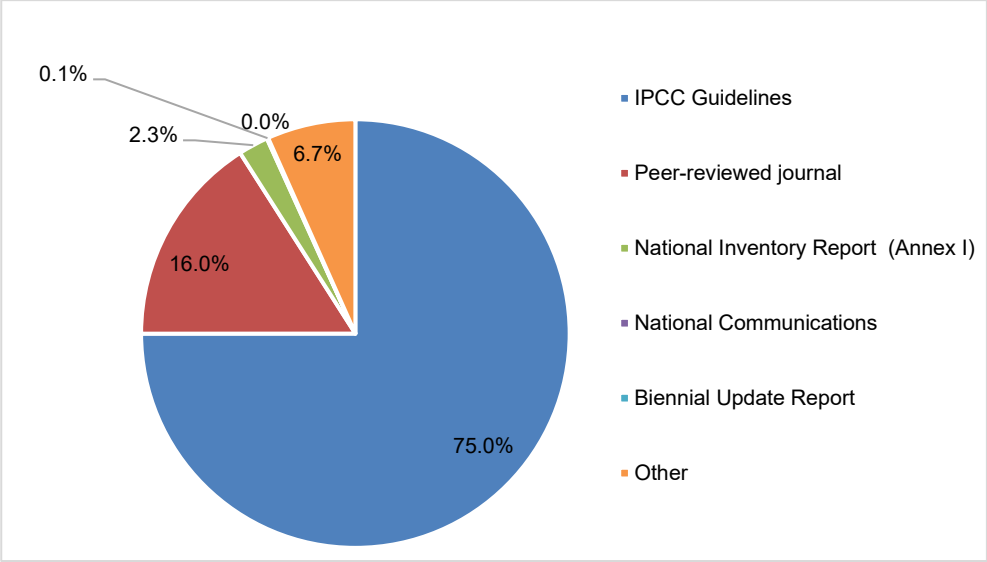
IPCC Emission Factor Database (EFDB): Preliminary analysis

Background

The overall objective of the IPCC EFDB is to be an always up-to-date companion for the IPCC Guidelines for National Greenhouse Gas Inventory (IPCC Guidelines) that is seen as a worldwide resource for greenhouse gas (GHG) inventory developers. The EFDB complements the information on emission factors (EFs) and other parameters of the IPCC Guidelines disseminating most recent scientific information on EFs and parameters.

The EFDB is meant to be a recognized data repository where users can find EFs and other parameters with background documentation or technical references. It contains default data from the IPCC Guidelines and data from other sources (e.g., peer-reviewed journal papers, national reports, etc.) including data for a source or sink not explicitly identified by the IPCC Guidelines.

Figure 1 EFDB data sources (as of 20 November 2024)



The EFDB is also one of the sources of data and information for authors in the production of the IPCC Guidelines. It is referred to as one of the sources of data information for inventory compilers in the 2006 IPCC Guidelines (Chapter 2, Volume 1 and Chapter 1, Volume 2) and 2019 Refinement (Chapter 2, Volume 1 and Chapter 4, Volume 2). It is also mentioned in the KP Supplement (Chapter 2).

The EFDB is open to any data proposals of EFs/parameters, and the data proposals submitted to the EFDB are assessed by the EFDB Editorial Board. So, it serves as a communication platform for experts and researchers to disseminate new EFs and other parameters in a timely manner to a worldwide audience of potential users.

The EFDB is recognized by Parties to the UNFCCC as a useful resource for inventory compilers (e.g., see FCCC/SBI/2011/5/Rev.1 and Decision 24/CP.19). It is considered useful particularly for developing countries and a training material on EFDB is included in the UNFCCC website (https://unfccc.int/process-and-meetings/bodies/constituted-bodies/consultative-group-of-experts/cge-training-materials/cge-training-materials-for-the-preparation-of-national-communications).

The EFDB operation and the work of its Editorial Board has seamlessly continued across the IPCC 4th, 5th and 6th assessment cycles.

EFDB Editorial Board

The Editorial Board comprises 36 members: 7 members for each of 5 sectors (Energy; Industrial Processes and Product Use; Agriculture; Land Use, Land-Use Change and Forestry; Waste) and a representative from Food and Agriculture Organization of the United Nations.

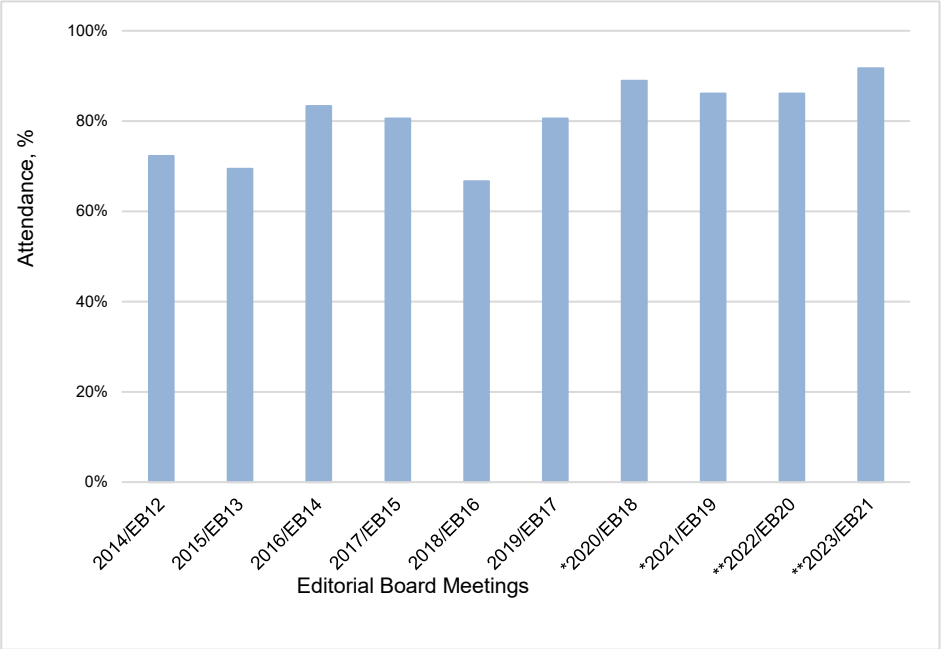
The term of the Editorial Board is two or four years. A key responsibility of the Editorial Board is to consider data proposals and decide whether proposed EFs or other parameters are acceptable or not according to the acceptance criteria: robustness[9], applicability and documentation (Roles and responsibilities of Editorial Board; Terms of Reference of the EFDB Editorial Board).

The aim of consideration/assessment of data proposals is not to undertake a full scientific assessment but to collect data that may be useful to potential EFDB users.

A meeting of the Editorial Board is organized annually and attended by the Editorial Board members (see Figure 2) to ensure consistency of the decision criteria for assessment of data proposals over time and between members of the Editorial Board.

[9] A robust emission factor or other parameter is one that, within the accepted uncertainty, is unlikely to change if there was repetition of the original measurement programme or modelling activity. The term “accepted uncertainty” refers to the uncertainty of IPCC defaults for that category.

Figure 2 Editorial Board Meetings



*Online; **Hybrid

EFDB Population

In general, the following approaches are used to populate the EFDB:

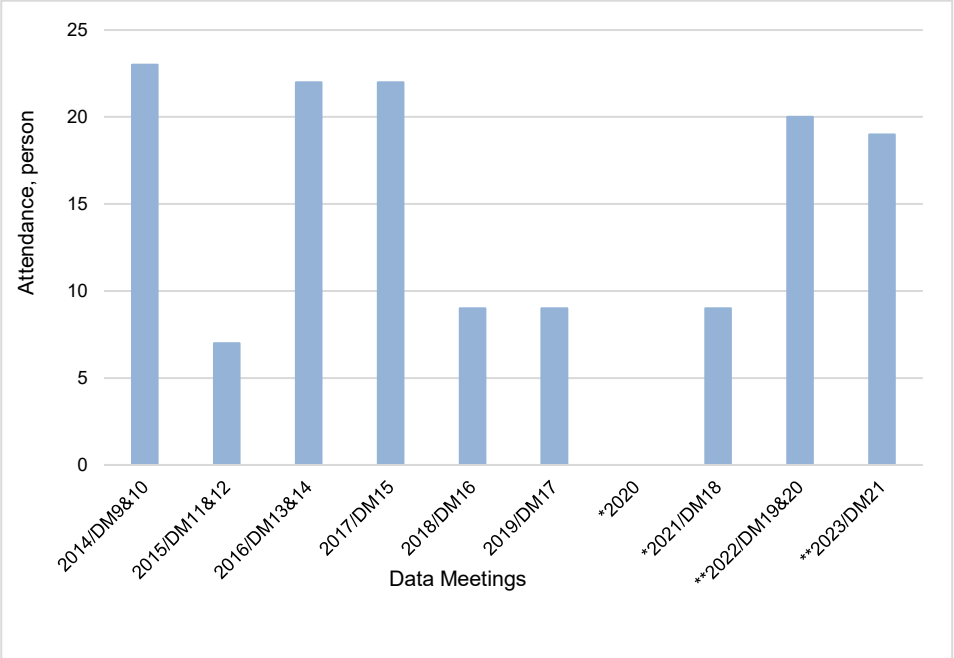
i. Voluntary submission of data proposals by experts.

ii. Literature search. The TSU conducts literature search and review to identify potential data/data sources and experts/data providers taking into account the areas identified in the technical assessment 10 of the IPCC Guidelines and prepares data proposals for consideration by the Editorial Board.

iii. Expert Meetings on Data for the EFDB (Data Meetings). The meetings have been organized since 2008 and are attended by both experts/data providers (see Figure 3) and the Editorial Board members of relevant sectors to consider data proposals submitted by experts/data providers for inclusion in the EFDB.

[10]Technical assessment of IPCC Guidelines carried out through an on-line questionnaire survey and four expert meetings in 2015 and 2016. The assessment identified a number of issues/areas that can be addressed by utilizing the EFDB.

Figure 3 Data Meetings

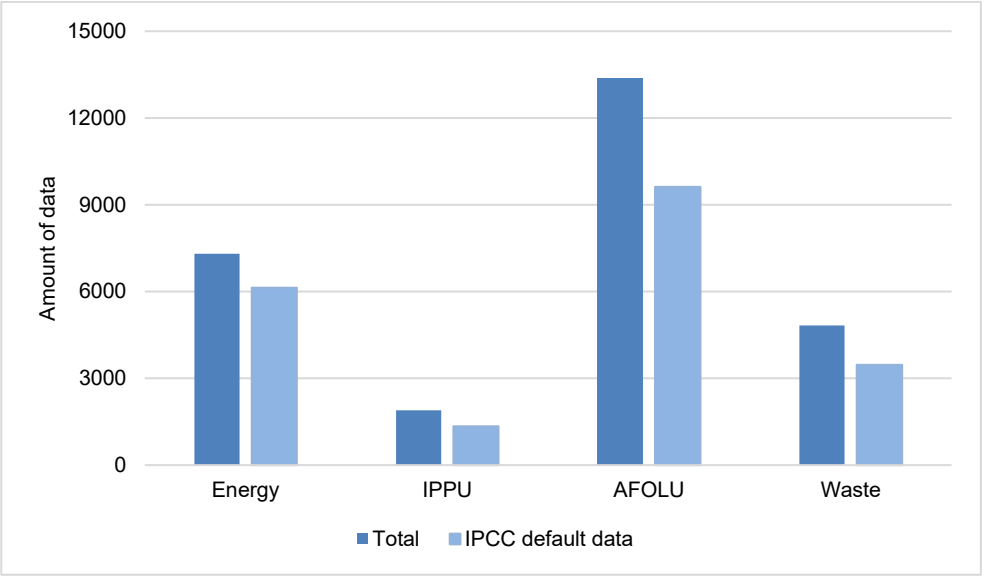


*Online; **Hybrid

The population of the EFDB has been increasing over the years. A total of 6263 data from sources other than IPCC Guidelines were accepted in 2014-2023 by Editorial Board and included in the EFDB. As of November 2024, the EFDB contained 27424 data (includes some default data from the 2019 Refinement): IPCC default data accounted for 75% and the remaining 25% was data from sources other than IPCC Guidelines (Figure 4). The uploading of the default data from the 2019 Refinement is underway.

Furthermore, the EFDB includes data on precursors (NH_3, CO, Nitrogen oxides ($NO+NO_2$), NMVOC and SO_2) from the *Revised 1996 IPCC Guidelines* and this could be expanded with updated data from the Methodology report on Short-lived Climate Forcers (SLCFs) including background data used in deriving default values on SLCFs.

Figure 4 EFDB data

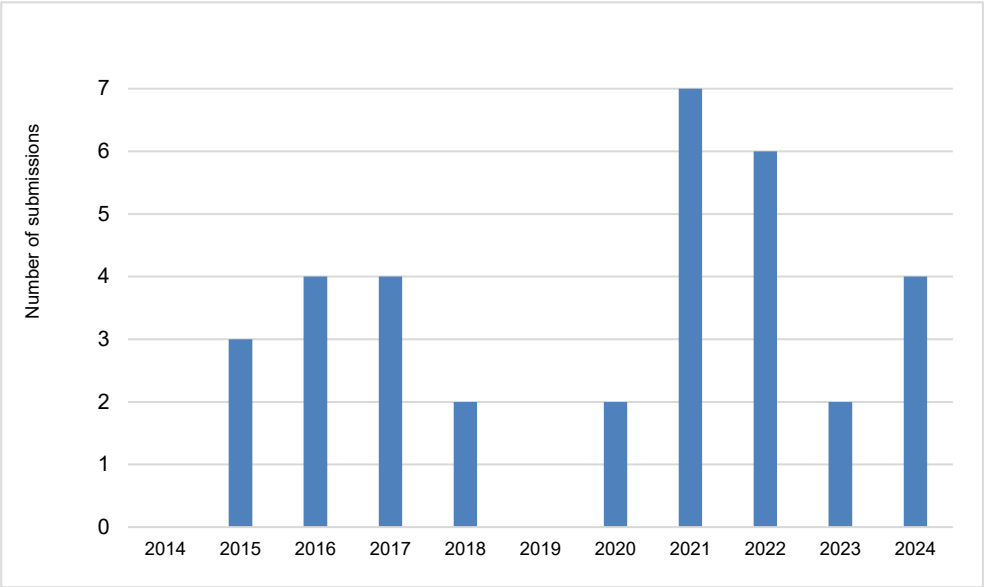


As at November 2024

However, there are some challenges in populating the EFDB.

- *Small number of voluntary submissions of data proposals* (see Figure 5). In 2013-2023, the number of voluntary submissions ranged from 0 to 7. Since the 21st Meeting of the Editorial Board (EB21) held in May 2023, TSU received 4 voluntary submissions: IPPU (1), AFOLU (2) and Waste (1). The data proposals for the Waste sector are considered by the Editorial Board through exchange of emails following the current procedures, and the data accepted by the Editorial Board are included in the EFDB. While the Editorial Board has responsibility for consideration and acceptance of data proposals, the TSU has overall responsibility for facilitating the data consideration process and ensuring that data is dealt with in a timely manner.

Figure 5 Voluntary submissions



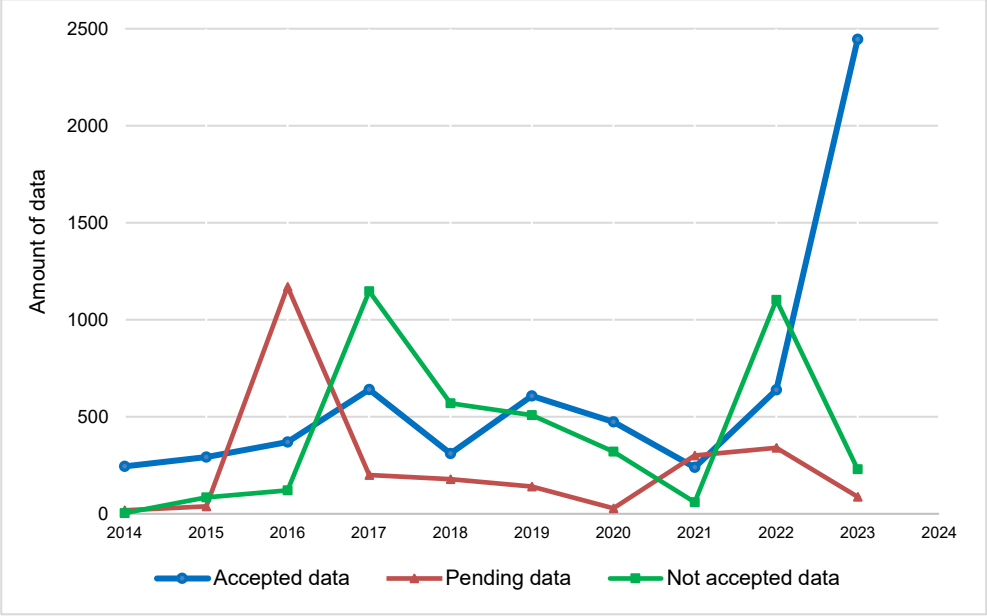
- *Literature search/review to identify potential data, data sources and data providers is carried out mostly by TSU.* A considerable amount of pending data. Figure 6 below shows amount of data accepted, not accepted and pending each year in 2014-2023. A large amount of pending data in 2016 comprises 1162 LULUCF sector data from other databases and 8 for Agriculture sector. Most of the data pending from 2016 was considered in 2017 by the Editorial Board at its EB15 but not accepted for inclusion into the EFDB. This resulted in a large amount of not accepted data in 2017 (Figure 6). A high amount of not accepted data in 2022 (EB20) is mostly LULUCF sector data (973 data which includes data pending from 2021). Furthermore, a large amount of data (1783 data) accepted in 2023 at EB21 includes data for Agriculture sector from other database/large datasets. However, it should be noted that in 2015 the Editorial Board agreed to provide at the EFDB website only links to other databases which contain suitable data considering that databases are dynamic, diverse and not all data in other databases might be suitable for the EFDB. It should also be noted that there were views among Editorial Board members that populating the EFDB with useful data is more important than only increasing the amount of data in the EFDB, and it should be useful as much as possible but not too big.
- *Assessment of data proposals by the Editorial Board occurs in recent years mostly at annual meetings of the Editorial Board*. However, according to the EFDB procedure, the work can be carried out through exchanges of e-mails (Section 3 of Roles and Responsibilities of Editorial Board; paragraphs 11-14 of Terms of Reference of Editorial Board).
- *Uploading the data to the EFDB*. The process involves more than just filling out data entry forms (Excel macro-enabled template). These forms include various fields and have specific technical requirements (e.g., mandatory fields, character limits in certain

fields, data format standards) which must be carefully adhered to. TSU prepare and conduct quality control of all data entry forms. It requires considerable amount of time particularly in the case of large datasets to ensure that data/information entered in data entry forms are free of error and data entry forms meet technical requirements for uploading the data.

EFDB Structure and Outputs

The EFDB contains data with background documentation in fields of each data. The output table of data search contains background information in data fields (columns) and the data field may contain several information e.g., “Region/Regional Conditions” field can contain country, region, climate zone etc. and users can narrow down the output table by defining and applying filters on the various columns/data fields of the output table. For example, if you are looking for data applicable to Asia, you can type the keyword “Asia” in the filter textbox above the column for “Region/Regional Conditions”. However, the structure of the EFDB (e.g., data fields) may need to be modified/changed to facilitate EFDB data search by users.

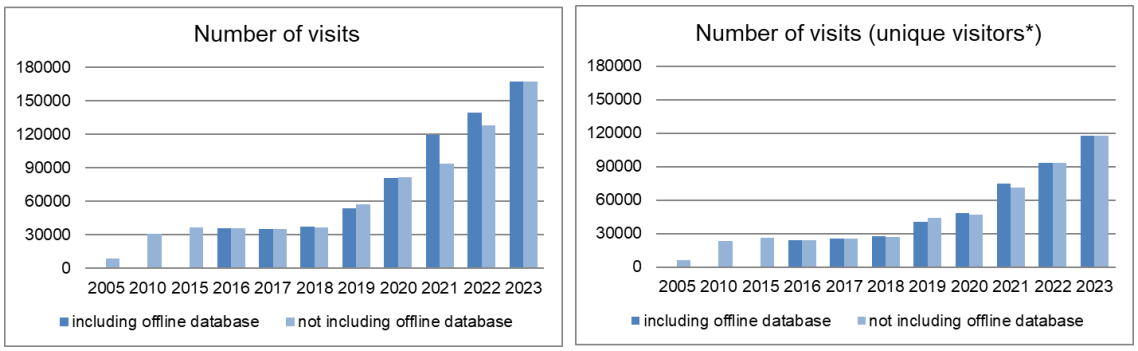
Figure 6 Data accepted, pending and not accepted



EFDB Usage

The number of visits to EFDB website has been increasing notably since 2018 (Figure 7).

Figure 7 Number of visits to EFDB



*A unique visitor is a user who has made at least 1 hit on 1 page of the website during a month. If the user makes several visits during the month, it is counted only once. The annual number of unique visitors count shows monthly totals.

The web application of the EFDB was upgraded with a new interface in 2018 and upgraded versions of the offline applications were released in 2018 and 2019.

Figure 8 EFDB download

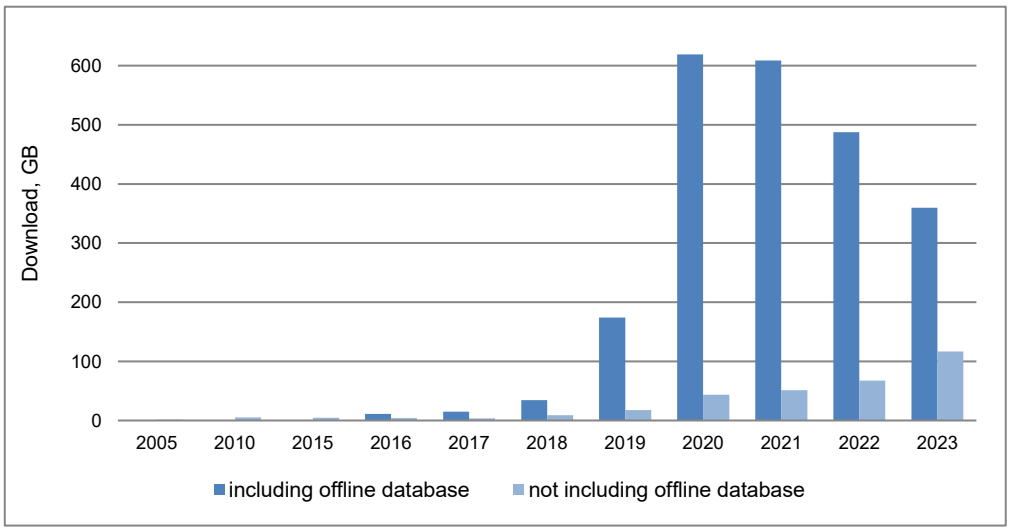
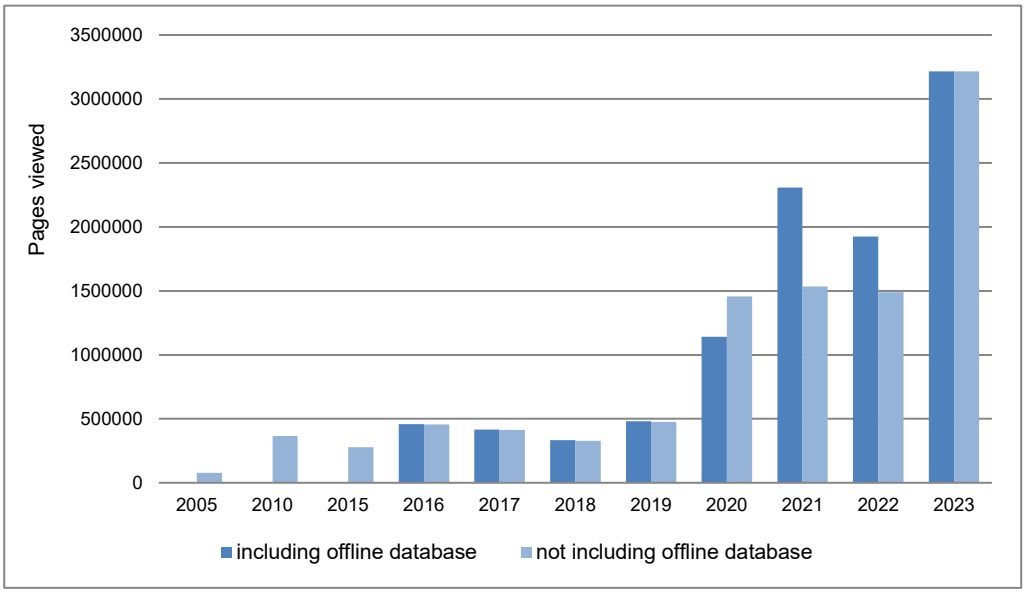


Figure 9 Pages viewed/accessed



The amount of data downloads and pages viewed/accessed have also been increasing. The offline application/database was removed in September 2023. In 2023, direct access to the EFDB accounted for 96.0% (96.1% in 2022) of total access.

To collect more information on the EFDB usage including information on who are users and data searched and respond to needs of users, a questionnaire for EFDB users has been prepared by TSU (not published).

Other Supporting Materials/Information for the EFDB assessment

Extra page

The Editorial Board considered that data which do not meet the acceptance criteria may still be useful for users (e.g., if the data are for categories/gases where there is a serious scarcity of data in the EFDB) and agreed to have an "extra" page on the EFDB website to provide such supporting information and data. It was highlighted that such an "extra page" is not meant to be a depository of data which was not accepted by Editorial Board. The "extra page" (https://www.ipcc-nggip.iges.or.jp/EFDB/otherdata.php) was set up in 2017 and currently includes 9 data sets: Energy (1), IPPU (7) and Waste (1).

Data from other Databases

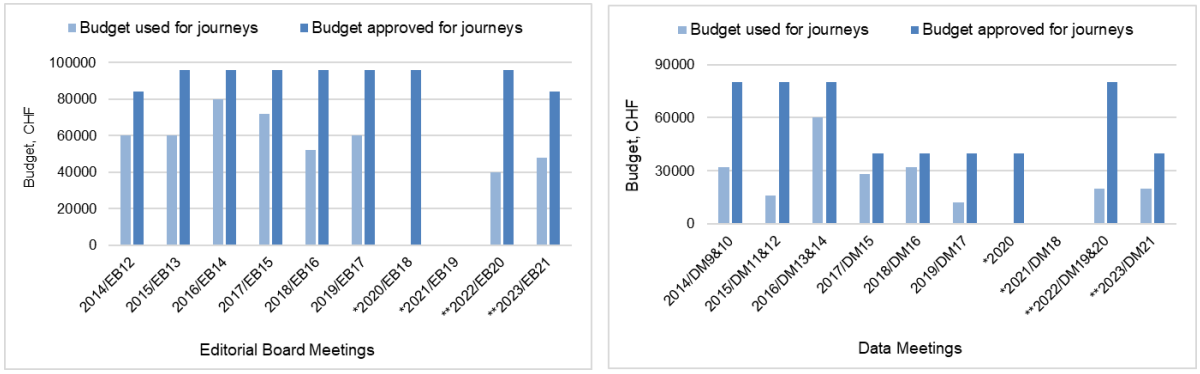
The Editorial Board, noting that databases are dynamic, diverse and not all data in other databases might be suitable for the EFDB, agreed to provide at the EFDB website only the links to other databases which contain suitable data. A webpage with links to such external databases was created on the EFDB website at https://www.ipcc-nggip.iges.or.jp/EFDB/otherdb.php.

Relevant Documents

- Terms of Reference (TOR) of the EFDB Editorial Board (EB21_Doc8_Updated TOR of EB.pdf)
- Role and responsibilities of Editorial Board (EB21_Doc4_Procedures.pdf)
- Proposal on how to deal with large data sets (EB21_Doc7_Proposal on dealing with large data sets.pdf) agreed at 15th Meeting of the Editorial Board.
- EFDB data entry form (EFDB_Bulk_Import_3.3.xlsm) available at https://www.ipcc-nggip.iges.or.jp/EFDB/main.php

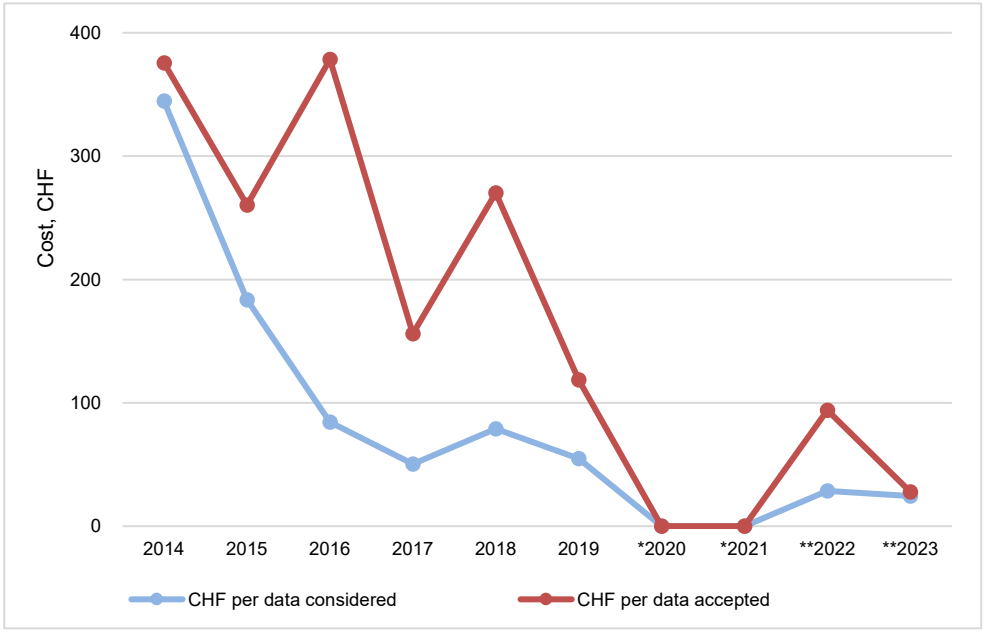
IPCC Trust Fund Budget

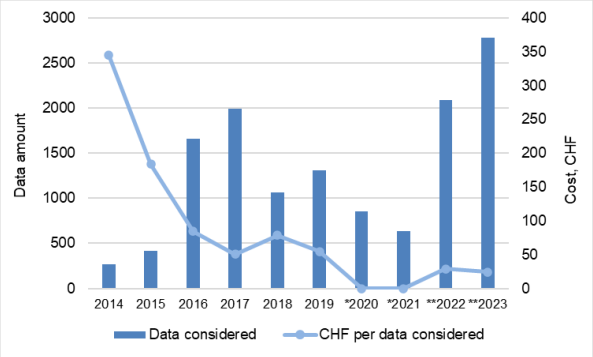
Figure 10 Budget for journeys of participants to attend the EFDB Meetings

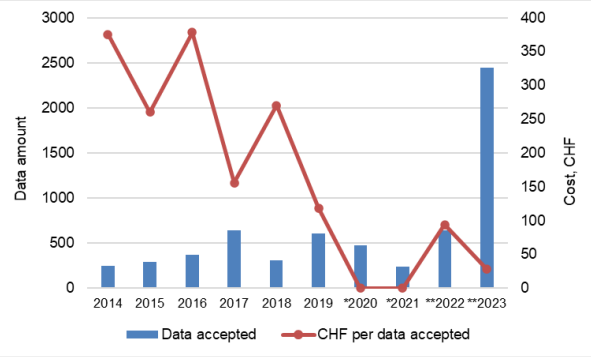


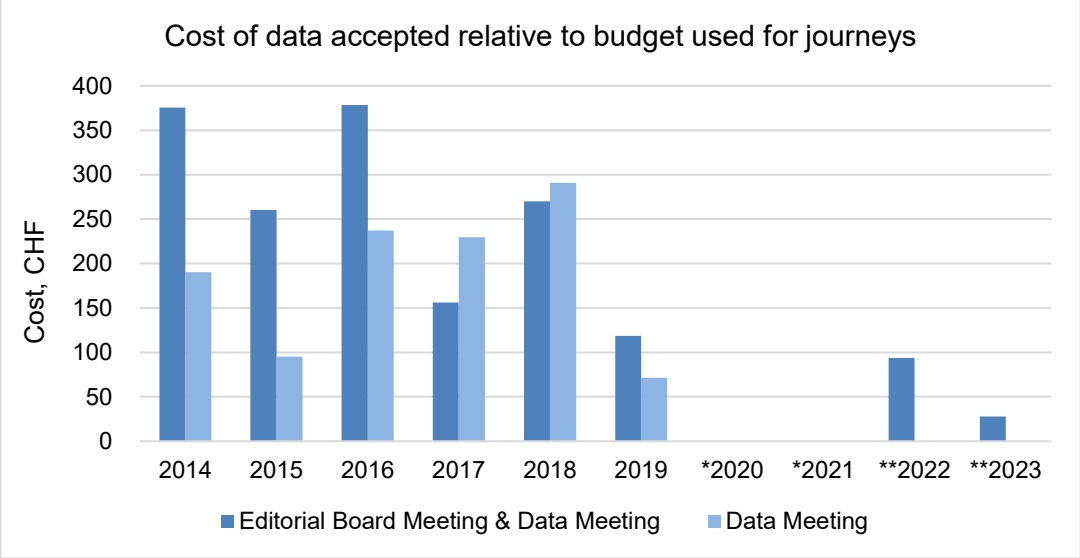
(*Online; **Hybrid)

Figure 11 Estimated cost of data relative to budget used for journeys









Editorial Board Meeting (EB) and Data Meeting (DM)	Cost of data accepted relative to budget used for journeys of participants to attend EB Meeting and Data Meeting	
	EB and DM	DM
2014 (EB12/DM9&DM10)	376	190
2015 (EB13/DM11&DM12)	260	95
2016 (EB14/DM13&DM14)	378	237
2017 (EB15/DM15)	156	230
2018 (EB16/DM16)	270	291
2019 (EB17/DM17)	119	71
*2020 (EB18/DM18)	0	No DM
*2021 (EB19/DM18)	0	0
**2022 (EB20/DM19&20)	94	
**2023 (EB21/DM21)		

*Online; **Hybrid

Summary

1. In general, the population of the EFDB has been increasing over the years. Although the amount of data from sources other than the IPCC Guidelines is increasing the amount is still low accounting for 25% of the total amount of data (includes some default data from 2019 Refinement, as of November 2024) in the EFDB.
2. Several challenges in enhancing the population of the EFDB:
    - Data proposals and data sources for consideration by the Editorial Board are prepared mostly by the TSU.
    - Experts/data providers for the Expert Meetings on Data for the EFDB are identified mostly by the TSU.
    - A low number of voluntary submissions of data proposals by outside experts/data providers
    - Assessment of data proposals occurs mostly at annual meetings of the Editorial Board
    - Considerable number of data pending from annual meetings
3. Improvement in procedures and criteria for assessment of data proposals including data from other databases and large datasets could help to enhance the population of the EFDB and overall performance.
4. The usage of the EFDB (e.g., number of visits, pages viewed, amount of data downloaded) has been increasing. However, an improvement in the structure of the database making it well-structured with clearly defined fields could further enhance the usefulness of the EFDB.

Appendix 1

Introduction

This annex aims to present some more detailed analysis on what is included in the EFDB and the parameter values. To allow for a more detailed analysis the work was limited to the energy sector, in particular stationary combustion, transport and fugitive emissions. Data in the EFDB for category 1.A.5 – Non-Specified and mobile off-road machinery included under category 1.A.4 was limited; analysis was not completed for these sectors. The analysis presented also excludes data in the EFDB from the IPCC guidelines (2006 Guidelines, 2019 Refinement and 1996 Guidelines).

The assessment focuses on three main areas. The first is providing a more detailed look at what is included in the EFDB, assessing the types of parameters as well as the number of data entries for information of interest such as fuel type or subcategory.

A potential use of the EFDB would be to house individual data (e.g. data from one site) and underlying data to national/regional averages. This analysis therefore also assesses how much of this individual data is already included.

The final assessment area is comparisons between the 2006 Guidelines T1 defaults and the non-default data included in the EFDB, where there is enough data available. Ideally, this would be completed only with individual data as this would be a more robust approach than comparing averages with averages. The ability to complete this analysis is however significantly limited by data availability, and the ability to identify individual data from the data fields in the EFDB.

Methodology

The EFDB data used for the assessment was downloaded 30th November 2024[11]. The analysis was completed using R, a programming language developed for statistical computing.

While much of the analysis could be completed using the exact values provided in the EFDB, there was some information that had to be standardized and pulled out of the description field:

- Parameter type (e.g. EF, NCV etc.): this information was determined using the description and units provided.
- Type of data[12] (e.g. regional/country averages or individual data): this information was determined from the description and 'Type of parameter' field. It was assumed that if

[11] The EFDB data was recently republished to include the 2019 Refinement data but the newly incorporated data would not impact the results of this work.
[12] The EFDB includes a data field 'Type of parameter' which is different from this 'Type of date' field added for this analysis.

the description did not explicitly state that the data was for one site (individual data point) then the data was an average for a region or country.

It is important to note parameters could be reported for multiple sectors, this means that the counts of data entries on a subcategory basis may be higher than the total number of data entries for a category. In addition, the count of data per category in this analysis will differ from that presented on the EFDB website. This is because this analysis looks solely at the reported sector for counts of more aggregated categories (e.g. 1.A) while the counts on the EFDB website sum the data for the category and all sub categories (e.g. the count for 1.A will sum all data reported with categories 1.A, 1.A.1, 1.A.2, 1.A.1.a, 1.A.1.a.i etc.).

Comparisons with the 2006 Guidelines default T1 factors use both histograms and kernel density curves [13], the latter of which can be considered smoothed histograms. The histograms use 10 'bins' and present the density[14]. The kernel density curves assume normal distribution and use a kernel bandwidth of 0.5[15], with the exception of road transport which used a bandwidth of 1. The analysis was only carried out where there were at least 10 data entries that could be compared with a T1 default parameter. This is likely too low for statistical significance or to draw any conclusions on the validity of the default data but does show the current spread of the data. This analysis was carried out on the individual data points and averages separately. Identifying if the data was from an individual site or a regional/country average was not straightforward and the information included in the EFDB export did not always make this information clear. Where it was not explicitly stated that the data was from one site, it was assumed that the data was a regional/country average. Ahead of the analysis some additional data manipulation had to be completed to convert the EFDB data to the units of the default parameters. Any units that could not be converted e.g. % for carbon content were excluded from the analysis.

Stationary Combustion

The following section provides analysis on the data included in the EFDB for stationary combustion, i.e. 1.A.1, 1.A.2 and 1.A.4 (excluding the mobile sources). General energy (1.A) was also included in this category and included, in addition to the expected calorific values and carbon contents, EFs. As the T1 default EFs in the IPCC guidance for stationary combustion are not category specific, the analysis has not been completed on a subcategory level.

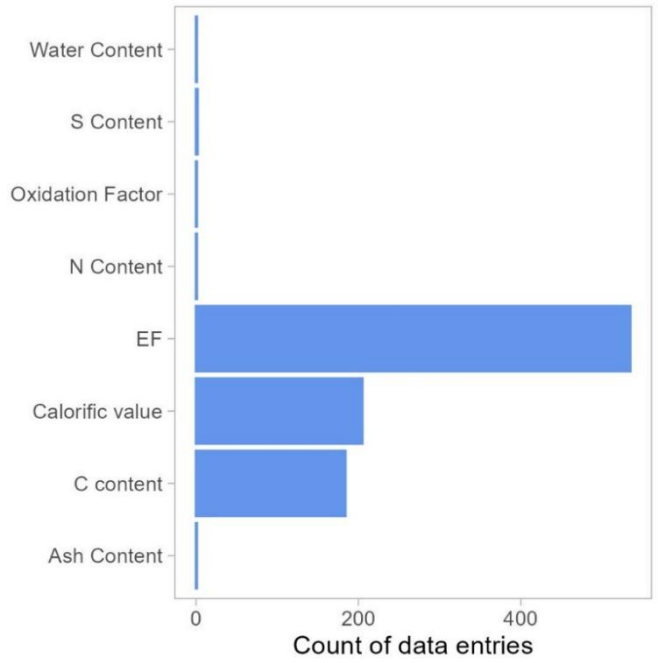
What data is included?

There are 930 data entries for stationary combustion and general energy. This is the highest amount compared to other energy categories. The majority of the data is emission factors (535), followed by calorific values and carbon contents (205, 184 respectively). One data

[13] Inbuilt into the ggplot2 package
[14] Frequency divided by bin width
[15] This value influences the smoothing and the value was selected based on the resulting charts, balancing the line smoothness and matching density values.

source also provided some parameters outside of the IPCC guidance (water, Sulphur, Nitrogen and ash content). There was also one data entry for oxidation factor.

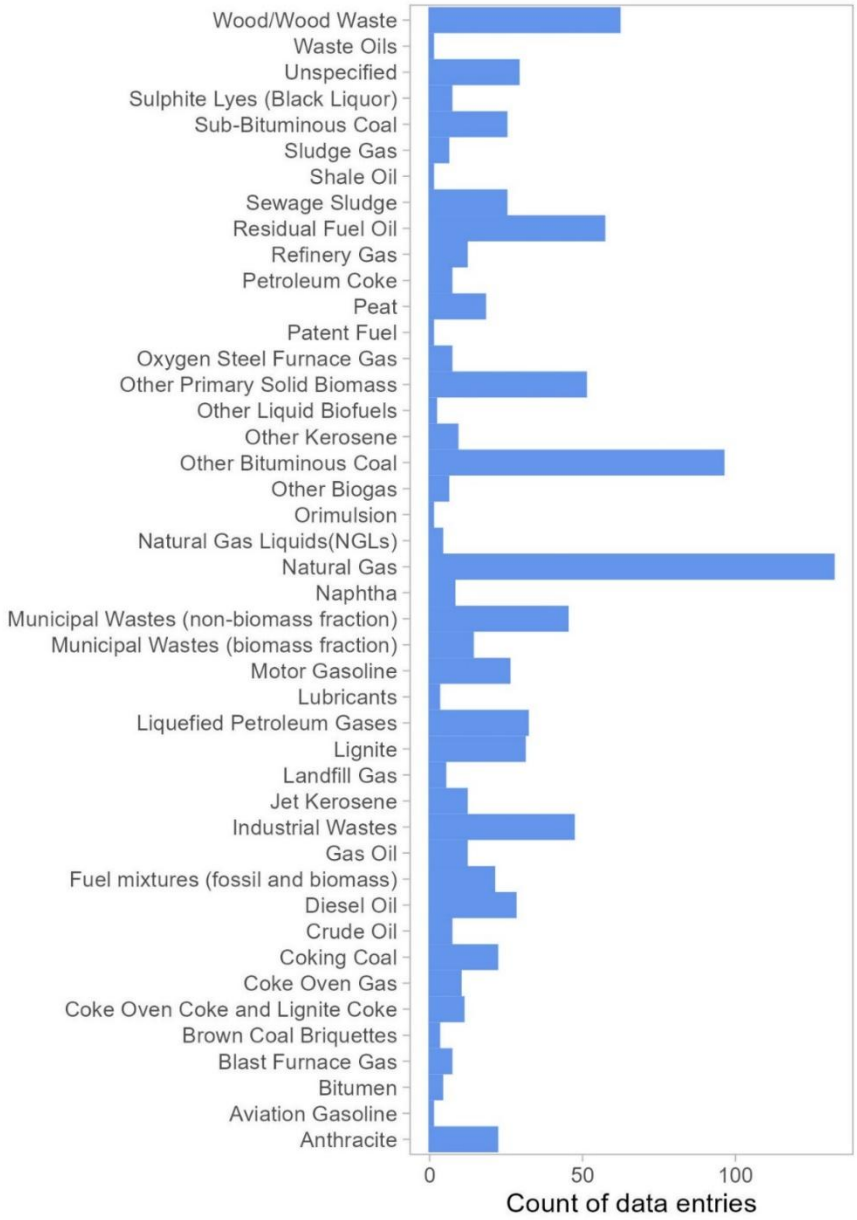
Figure 12: Number of data entries for stationary combustion by parameter type



The most common fuel for stationary combustion is natural gas (132) followed by other bituminous coal (96)*As* noted above these figures differ from those presented on the EFDB website due to the different method.

Figure 13. The most common sector was general fuel combustion activities (1.A), 49%, followed by energy industries and its subcategories (1.A.1, 1.A.1.a, 1.A.1.a.i etc.), 28%. As noted above these figures differ from those presented on the EFDB website due to the different method.

Figure 13: Number of data entries by fuel type for stationary combustion



In addition to parameters not included in the IPCC guidelines, there is a fair amount of data in the EFDB for this sector that is provided in different units to those in the guidelines (and not convertible). Examples include: emission factors on a GCV basis, % for carbon content, emission factors based on production quantities and NCVs for natural gas on a volume basis.

For stationary combustion the gas with the most parameters is CO_2, 66%. This is however slightly biased as both calorific values and carbon content are assigned to the gas CO_2. When just comparing EFs the split is much more even, with EFs from CO_2 contributing 42%, CH_4 28% and N_2O 29% to the total number of EFs. There are additionally two EFs each for the precursor gases NO_X and SO_X in the EFDB for stationary combustion.

Despite the additional data there is not complete coverage of all expected parameter, fuel and gas[16] combinations. Of the 162 EF combinations (by fuel and gas) with T1 default EFs in the 2006 Guidelines only 52% have an additional value in the EFDB.

As far as it was possible to identify, no individual site data has been included in the EFDB for stationary combustion. The data included are country specific or technology specific parameters from NCs, NIRs, peer-reviewed journals and other reports. The majority of the data is however from measurements (80%).

Comparison with the 2006 Guidelines T1 defaults

As it is concluded that no individual site data has been included for stationary combustion this analysis has been completed using the regional/national averages data, Figure 3, which was possible for nine emission factors for coking coal, natural gas, other bituminous coal and residual fuel oil. The blue bars are the histograms showing the density[17] and the solid red curve is the kernel density curve. The vertical dashed lines show the IPCC 2006 guidelines T1 default as well as the upper and lower limit. As this analysis was done on the regional/national averages no conclusions on the suitability of the default EFs can be drawn despite there being a high number of data entries for comparison with the natural gas EFs (45 for N_2O, 42 for CH_4 and 26 for CO_2). The remaining fuel, gas combinations had 20 or less data entries for comparison.

The distribution shows that, for the data analysed, the EFs in the EFDB are mostly within the upper and lower bounds of the default emission factors. This is however biased by the selection process, as included above. There are however some fairly significant outliers, the EFs for CH_4 from natural gas combustion having the most. All but two of the outliers presented in Figure 3 (EFs for CH_4 from natural gas which have no source provided) are included in the EFDB from Byung Moo Min and et al. (2008)[18] and the Finnish NIR (2008)[19].

[16] CO_2, CH_4 and N_2O

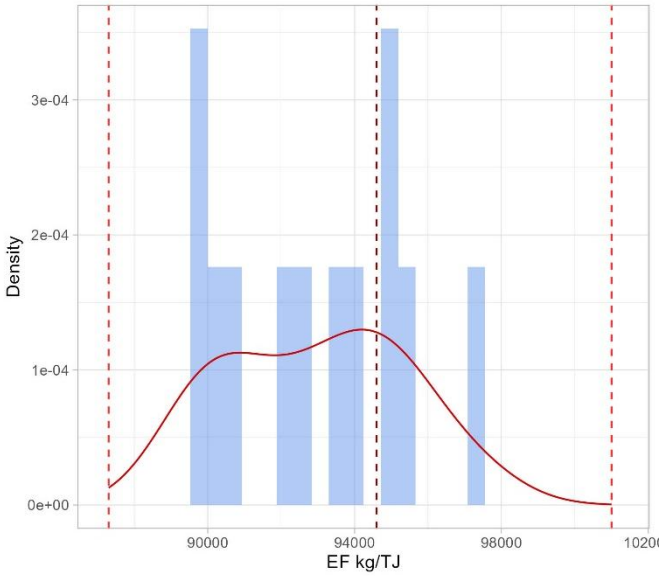
[17] Frequency divided by bin width

[18] Byung Moo Min and et al., Development of Country Specified Greenhouse Gases emission Factor(I) (2008), KEMCO-2007-31

[19] Finnish NIR 2008. Expert estimates by Statistics Finland based mainly on the VTT studies (Tsupari et. al., 2005, Tsupari et. al., 2006 and Tsupari et al., 2007); http://www.vtt.fi/inf/pdf/tiedotteet/2005/T2321.pdf

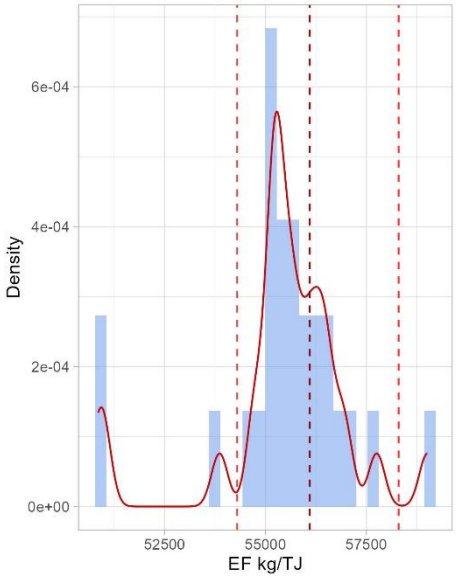
Figure 14: Comparison between EFDB data and 2006 Guidelines T1 defaults

EF for CO_2 from coking coal combustion



Mean: 92961

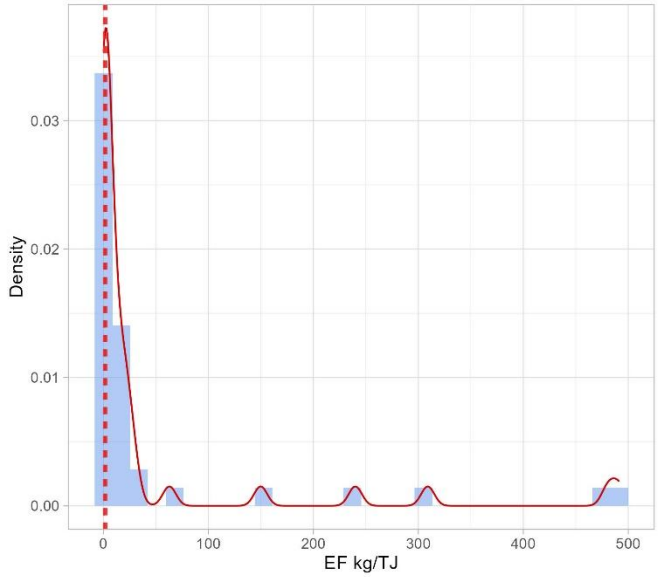
EF for CO_2 from natural gas combustion



Mean: 55489

Note: several EFs in the EFDB for this had a description of "carbon emission factor" but a unit of g/MJ. It was assumed that the EF was

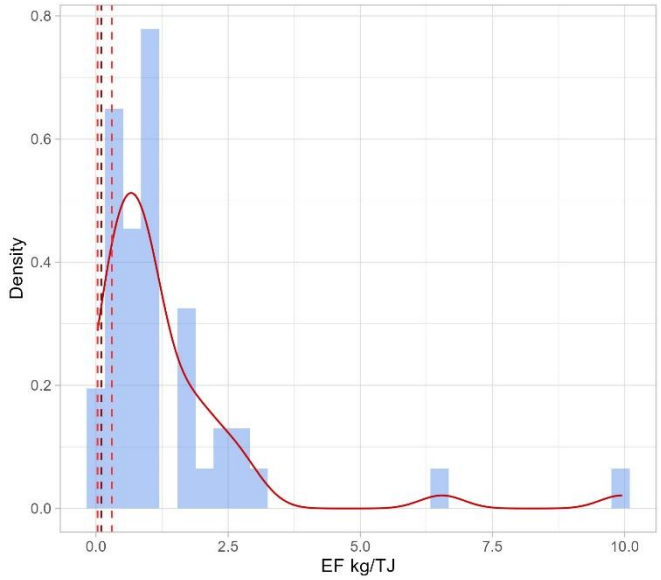
EF for CH_4 from natural gas combustion



IPCC 2006 GL T1 Default
IPCC 2006 GL T1 Upper and Lower limit

Mean: 47 kg/TJ

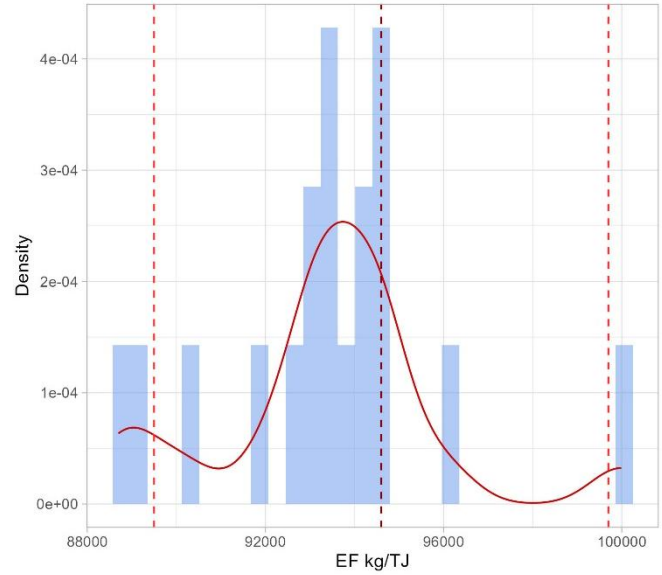
EF for N_2O from natural gas combustion



IPCC 2006 GL T1 Default
IPCC 2006 GL T1 Upper and Lower limit

Mean: 1.35 kg/TJ

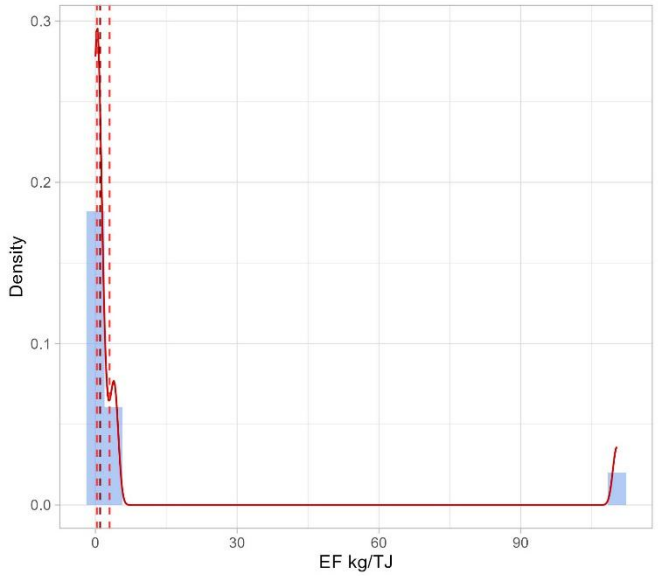
EF for CO_2 from other bituminous coal



IPCC 2006 GL T1 Default
IPCC 2006 GL T1 Upper and Lower limit

Mean: 93422

EF for CH_4 from other bituminous coal

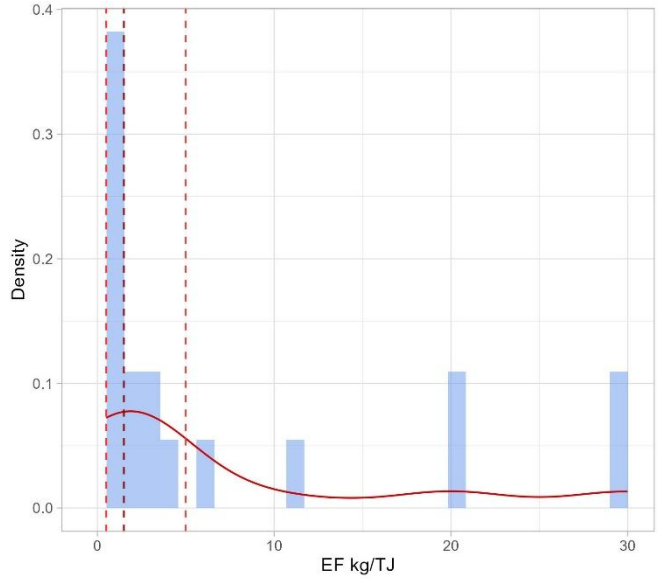


IPCC 2006 GL T1 Default

IPCC 2006 GL T1 Upper and Lower limit

Mean: 9.5 kg/TJ

EF for N_2O from other bituminous coal

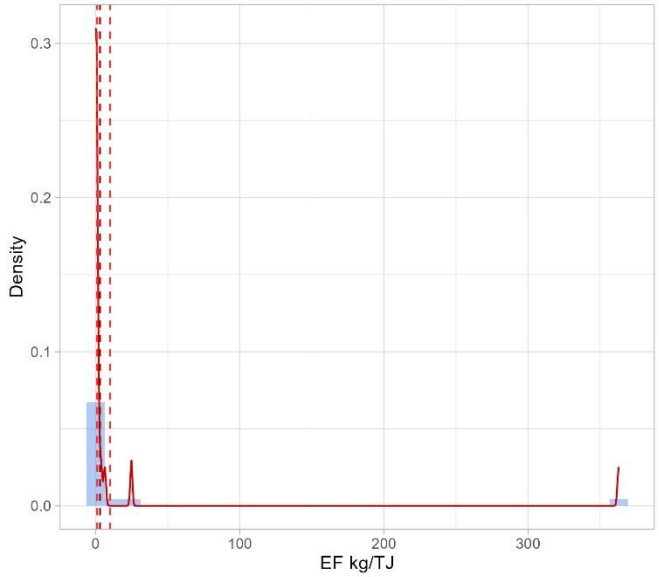


IPCC 2006 GL T1 Default

IPCC 2006 GL T1 Upper and Lower limit

Mean: 7.7 kg/TJ

EF for CH_4 from residual fuel oil

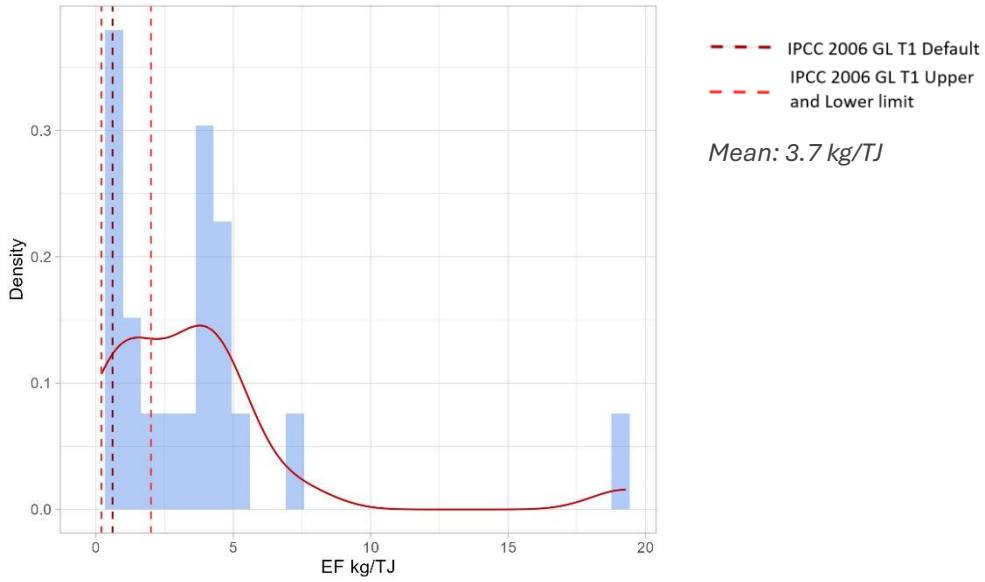


IPCC 2006 GL T1 Default

IPCC 2006 GL T1 Upper and Lower limit

Mean: 21 kg/TJ

EF for N_2O from residual fuel oil



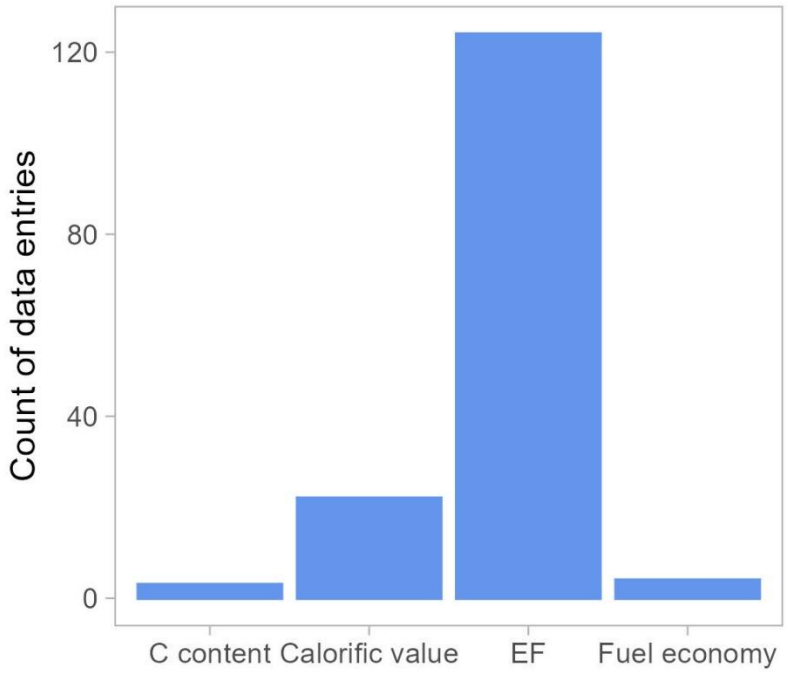
Transport

The following section looks at parameters in the EFDB for transport categories (1.A.3 and subcategories). While the IPCC guidelines default T1 CO_2 EFs for transport are not subcategory specific, the CH_4 and N_2O ones are. This analysis has therefore been completed considering the IPCC categories of the data.

What data is included?

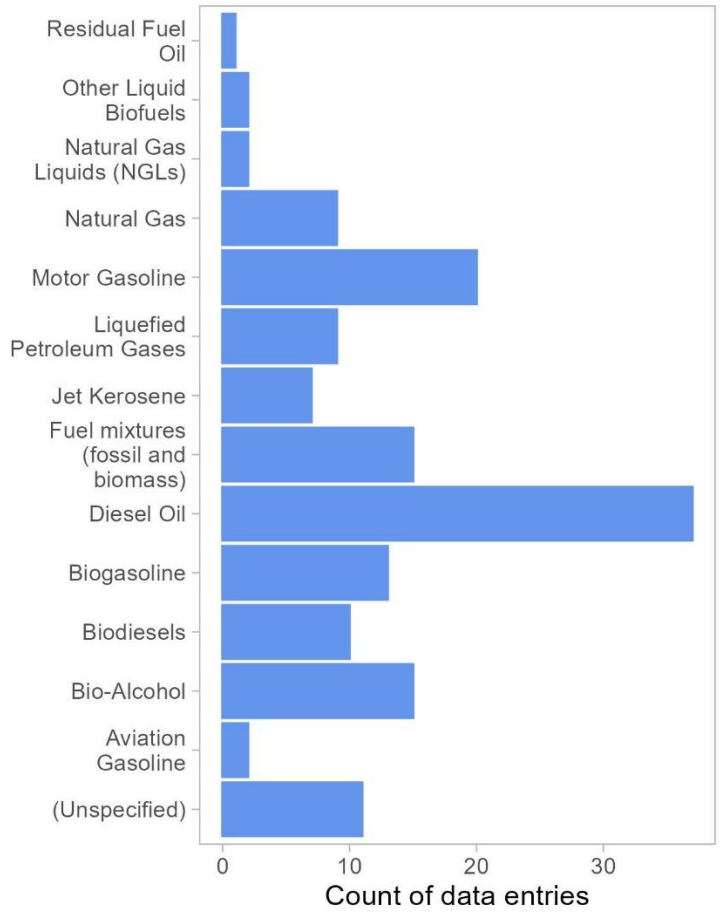
There are 153 data entries in the EFDB for transport, excluding those from the IPCC guidelines, although many are applicable to multiple categories. As Figure 15 shows these are predominantly EFs, with a small portion of calorific values (NCV and GCV) and one carbon content. There is also one fuel economy value, which is beyond the information provided in the IPCC guidelines but could be useful for inventory compilers.

Figure 15: Transport EFDB data by parameter type



The most common fuel these parameters are provided for is diesel oil with motor gasoline in second, Figure 16. This is likely due to the high proportion of the data being for the road transport sectors (78%).

Figure 16: Number of data entries by fuel type



EFs for N_2O (34%) were the most populous, although the number of EFs for CH_4 and CO_2 were close in number (27% and 25% respectively). There were also eight EFs for the precursor gases CO and NO_X.

The IPCC guidance for transport emissions provides default data on an energy basis, although there is some example data which is provided on a distance basis from the US and Europe. As for stationary combustion, for transport the EFDB includes data in different units to what is included in the IPCC guidelines. Examples include, emissions per mass of fuel, emissions per landing or take-off, NCVs on a volume basis and emissions per distance for additional countries.

Similar to stationary combustion there is not complete coverage for all EF, fuel, gas[20] combinations included in the 2006 Guidelines for this category. Of the 50 EF combinations (by sector, fuel and gas) with T1 default EFs only 38% have a non-guidelines value in the EFDB. There are five fuels, excluding unspecified, for which non-default data are included in the EFDB that are not included in the T1 EFs of the transport chapter. However, T1 EFs included in the transport chapter of the 2006 Guidelines are often a sub-

[20] CO_2, CH_4 and N_2O

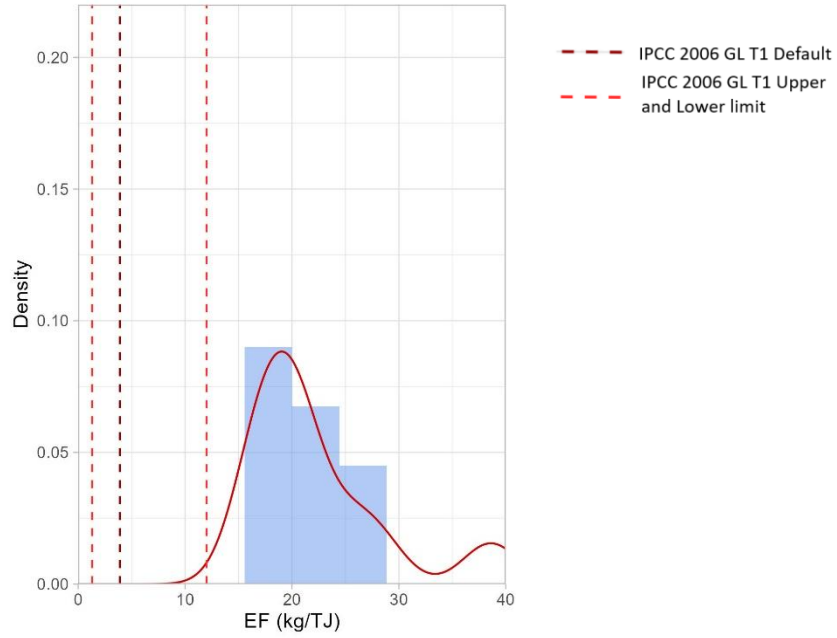
set of those included in the stationary combustion chapter. The EFs for these five fuels are therefore not considered additional information.

Unlike for stationary combustion and fugitive emissions, single points of data are harder to define. The EFDB does not include any data for individual vehicle engine types, as far as can be identified, the most disaggregated are emission factors for engine classes, EURO standards with specified abatement technology.

Comparison with the 2006 Guidelines T1 defaults

As no single points of data could be identified for transport this analysis has been completed using the regional/national averages data. Due to data availability this was only possible for the EF for N_2O emissions from 1.A.3.b.iii - Heavy-duty trucks and buses, Figure 17. All of the EFs are above the upper limit of the 2006 Guidelines T1 default (12 kg/TJ), and many are more than double this upper limit. However, as these are regional or country averages and only ten data entries were suitable for this comparison no conclusions can be drawn on the suitability of the default emission factor.

Figure 17: Assessment of the spread of N_2O EFs for 1.A.3.b.iii - Heavy-duty trucks and buses



Fugitive emissions

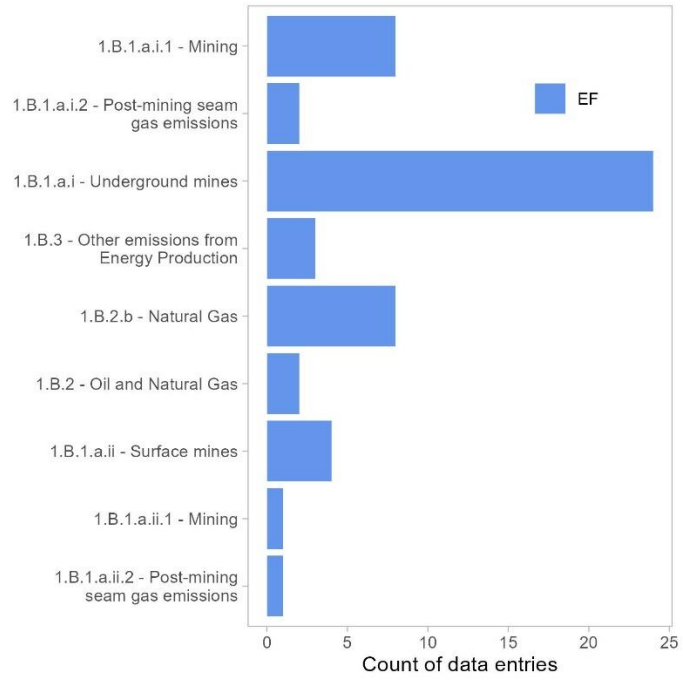
The following section looks at parameters in the EFDB for fugitive emissions from energy production (1.B and subcategories). EFs for this category are subcategory specific and therefore the analysis has been completed considering the IPCC categories of the data.

What data is included?

Overall there were 51 data entries for fugitive emissions, excluding values from the IPCC guidance. As Figure 7 shows all these are EFs. The category with the highest number of

data entries was 1.B.1.a.i – Underground mines, by a significant margin, making up 45% of the fugitive emissions data. The remaining categories made up 15% or less each. Overall, there are significantly more EFs for coal mining than oil and natural gas systems included in the EFDB, with only five EFs for natural gas systems and none for oil. The three EFs for 1.B.3 – Other emissions from Energy Production is for geothermal energy production.

Figure 18: Count of EFDB data entries per sector and parameter type



The majority of the EFs included in the EFDB for fugitive emissions were for methane which is expected as this is the most significant GHG emitted from coal mining, which has the most coverage in the EFDB. There are no default EFs for CO_2 from coal mining making the EFDB a potentially useful resource for this information, although only seven data entries have been included so far.

Compared to stationary combustion and transport, there is relatively little data provided in units that differ from those included in the IPCC guidance. All the EFs provided for mining activities are in m^3/tonne matching the IPCC guidance. However, all the EFs for emissions from natural gas systems differed from what was provided in the guidance including emissions per event (well completion flowback and unloading) and emissions per well.

Fugitive emissions has the lowest coverage for all EF, gas[21] combinations from the 2006 Guidelines of the categories analysed. Of the 18 EF combinations (by sector, fuel and gas) with T1 default EFs only 22% have a non-guidelines value in the EFDB. Much of the data in the EFDB for this category is provided at a more aggregated category (e.g. 1.B.1.a.i - Underground mines compared to 1.B.1.a.i.1 - Mining and 1.B.1.a.i.2 - Post-mining seam gas emissions) which is contributing slightly to this low coverage, but not significantly.

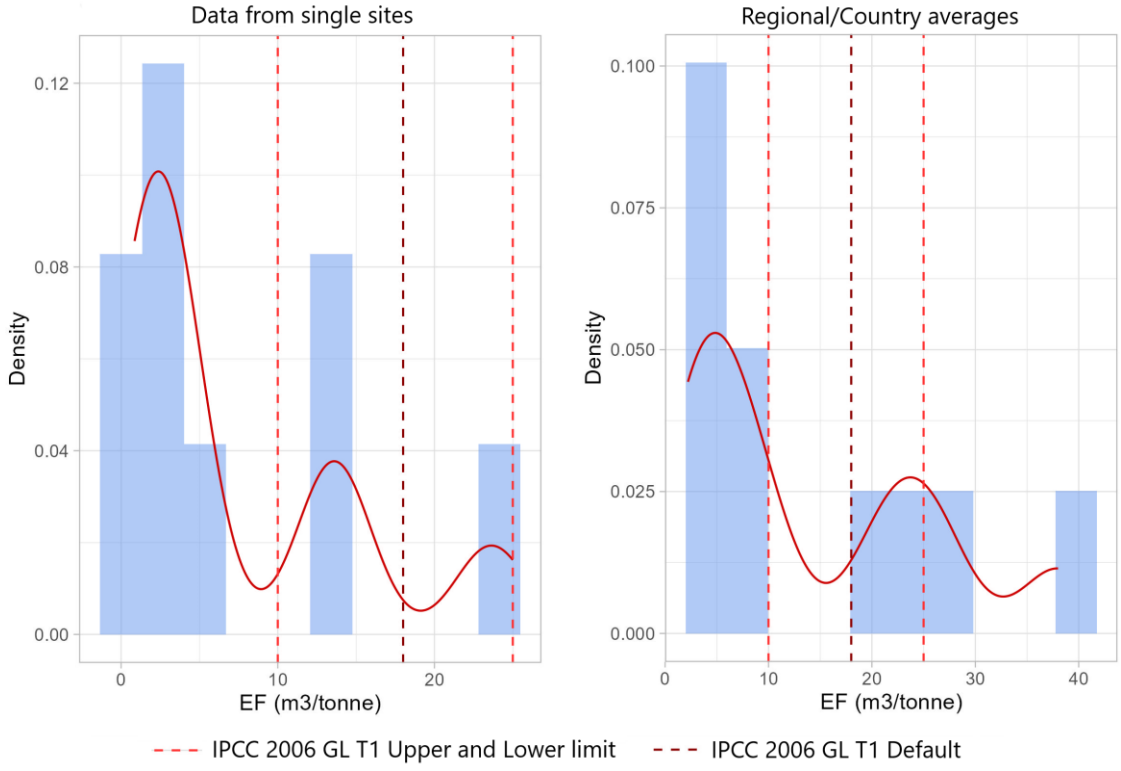
The majority of the fugitive emissions data was average values from measured data sourced from NC and NIR as well as other reports and peer-reviewed journals. There were however some single measured data, ten for CH_4 emissions from 1.B.1.a.i - Underground mines and two for CO_2 from the same subcategory.

[21] CO_2, CH_4 and N_2O

Comparison with the 2006 Guidelines T1 defaults

There was only one sector with any single data points in the EFDB for a comparison of the spread of data values with the default EFs from the 2006 Guidelines (for methane), 1.B.1.a.i - Underground mines, which based on the provided description were all identified as active mining. The comparison with the 2006 Guidelines T1 default EF was therefore completed for this sector only for the single data points and regional/national averages separately, Figure 19. While there are some EFs that are higher or lower than the upper and lower limits of the 2006 Guidelines T1 default, none of the data are significant outliers for either the data for single sites or regional/country averages. The mean of the single site data is 14.1 m^3/tonne which is between the upper and lower bound of the default EF (which itself is 18 m^3/tonne). The mean of the regional/country averages is 7.3 m^3/tonne which is below the lower limit of the default EF. There is quite high proportion of the data that are below the lower limit of the default EF in both cases but there is not enough data to draw any conclusions on the accuracy of the default EF.

Figure 19: Assessment of the spread of methane EFs for 1.B.1.a.i - Underground mines



Conclusions

Analysis was done on the energy sector data in the EFDB, excluding data from the IPCC guidelines (2006 Guidelines, 1996 Guidelines and 2019 Refinement). Data in the EFDB for category 1.A.5 – Non-Specified and mobile off-road machinery included under category 1.A.4 was limited so the analysis focused on 1.A.1 – Energy Industries, 1.A.2 -

Manufacturing Industries and Construction, 1.A.4 – Other Sectors (stationary combustion), 1.A.3 - Transport and 1.B Fugitive Emissions.

There is not a huge amount of data, once that from the IPCC guidelines has been removed, and there are significant gaps in the coverages of gases, sectors and fuels. When comparing the potential combinations of categories, fuels and gases for EFs, based on the EFs provided in the 2006 Guidelines it was found that only 52% of the default EFs had non-default data included in the EFDB. This was even lower for transport and fugitive emissions with 38% and 22% coverage respectively.

While housing individual measurement data is considered a potential use of the EFDB there is currently very little of this data for the Energy Sector. Of the data analysed only ten data entries could be identified as being measurement data for a single site, all in the fugitive emission category.

Comparison of the non-default data in the EFDB with the default parameters was limited by the amount of data available and very few categories or fuels had enough data points for comparison. This combined with the fact that, as far as it was possible to identify, the majority of the data was regional or country averages means that the data in the EFDB cannot be used to assess the validity of the default parameters. Regardless comparisons, where possible, showed that a significant portion of the non-default data was within the upper and lower limits of the T1 defaults. Even those outside of the limits were usually not significant outliers. This however is biased by the data selection process. The category with the most outliers was stationary combustion, which had some fairly significant outliers for CH_4 and N_2O EFs from a few fuels.

Reflections on searching the database

Completing the analysis on the data has highlighted some potential barriers to the usability of the database. The following observations are made:

1. The EFDB contains significantly more data than is available to export in bulk, this data is only available to download per data point.
2. Currently what the data is referring to (EF, NCV etc.) is included in the description and therefore there is no consistency. For example, and EF can be listed as EF, Emission Factor, emission factor or just emission. This could make it hard for users to search the database and is likely to be even more difficult for other sectors with more parameters.
3. The reference field is not always completed, potentially reducing trust of users in the data.
4. It is not always easy to find the document using the provided reference online, especially older papers. There is a non-mandatory field to provide the link but website links can change over time and this data is not available in the bulk export.

5. Whether data from the guidelines is a T1 or higher tier parameter is only sometimes included in the description.