UNCERTAINTIES IN THE EDGAR EMISSION INVENTORY OF GREENHOUSE GASES

Efsio Solazzo¹, Monica Crippa¹, Diego Guizzardi¹, Marilena Muntean¹, Marigita Choulga², Greet Janssens-Maenhout¹

¹ European Commission, Joint Research Centre, IPRC Ispra (VA), 21027, Italy
² European Centre for Medium Range Weather Forecasting, ECMWF, Reading, RG2 9AX, United Kingdom

ABSTRACT

The Emissions Database for Global Atmospheric Research (EDGAR) estimates the human-induced emission rates on Earth. EDGAR collaborates with atmospheric modelling activities and aids policy in the design of mitigation strategies and in evaluating their effectiveness. In these applications, the uncertainty estimate is an essential component, as it quantifies the accuracy and qualifies the level of confidence in the emission.

This study complements the EDGAR’s emissions inventory with estimation of the structural uncertainty stemming from its base components (activity data statistics (AD) and emission factors (EF), by (i) associating uncertainty to each AD and EF characterizing the emissions of the three main greenhouse gases (GHGs), namely carbon dioxide (CO₂), methane (CH₄) and nitrous oxide (N₂O), (ii) combining them, and (iii) making assumptions for the cross-country uncertainty aggregation of source categories.

It was deemed a natural choice to obtain the uncertainties in EFs and AD from the Intergovernmental Panel on Climate Change (IPCC) guidelines issued in 2006 (with a few exceptions), since the EF and AD sources and methodological aspects used by EDGAR have been built over the years based on the IPCC recommendations, which assured consistency in time and comparability across countries. While on one side the homogeneity of the method is one of the key strengths of EDGAR, on the other side it facilitates the propagation of uncertainties when similar emission sources are aggregated. For this reason, this study aims primarily at addressing the aggregation of uncertainties sectorial emissions across GHGs and countries.

Globally, we find that the anthropogenic emissions covered by EDGAR of the combined three main GHGs for the year 2015 are accurate within an interval of -15% to +20% (defining the 95% confidence of a log-normal distribution). The most uncertain emissions are those related to N₂O from waste and agriculture, while CO₂ emissions, although responsible for 74% of the total GHG emissions, account for approximately 11% of global uncertainty share. Sensitivity to methodological choices is also discussed.

INTRODUCTION

According to the latest release of the Emissions Database of Global Atmospheric Research (EDGAR version 5, Crippa et al., 2019, Crippa et al., 2020a), in the year 2015 the global greenhouse gas (GHG) emissions of carbon dioxide (CO₂), methane (CH₄) and nitrous oxide (N₂O) due to anthropogenic activities summed up to 48.1 Gt CO₂eq. CO₂ equivalent emissions (CO₂eq) are computed using the Global Warming Potential values from the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC). In the same year, the share of global CO₂eq from non-CO₂ GHG emissions (i.e., CH₄ and N₂O) was approximately a quarter. Measures put in place to attenuate temperature rise and to mitigate climate dynamics long-term changes have contributed to uphold the
role of CH₄ and N₂O. Their high warming potential compared to CO₂ and relatively shorter life-time (on average CH₄ persists in the atmosphere for approximately a decade, N₂O for over a century and CO₂ for more than 1000 years (NCR, 2010; Ciais et al., 2013)) allow to get on shifting from energy-related CO₂ to other, more readily responsive emission sources (Janssens-Maenhout et al., 2019, United Nations Environment Programme, 2019). At the same time, while for fossil fuel CO₂ emissions the uncertainty is relatively small and, overall, well defined, for CH₄ and N₂O the emission estimates are significantly more uncertain. In turn, emission reduction measures issued by national plans highly depend on the degree of uncertainty of sectors that are supposed to contribute to reach the desired reduction target. As depicted in the example by Olivier (1998) a sector contributing by 10% to the national reduction target may contribute to 5% or 15% if that sector’s emission factor is ±50% uncertain.

EDGAR aims to consolidate its position in supporting research and novel data/approach implementation in operational modelling, as well as becoming an independent tool supporting policy makers in monitoring and mitigation strategies. Therefore, a reliable quantification of the uncertainties should have the same degree of importance as the consistency and comparability of the emissions. This study evolves in this direction, by adding the uncertainty dimension to the EDGAR database, thus enhancing its value with much needed information on reliability and promote comparability with other datasets. Uncertainty reports are relevant, among other applications, for:

- scientific purposes, e.g. assessing robustness of long-term emission trends, or provide a-priori state of comparison with independent top-down estimates (Bergamaschi et al., 2018), or aid in network design (Super et al., 2020);
- inter-comparison studies (Choulga et al., 2020; Petrescu et al., 2020);
- assessing the feasible potential of mitigation strategies (e.g. Van Dingenen et al., 2017).

This study adds the uncertainty component to the EDGAR data by devising methods to propagate the uncertainty introduced by activity data (AD) and emission factors (EFs) to any combination/aggregation of sources, countries, and GHGs. Methods, aggregation strategies and dependencies are presented and investigated. Analyses are conducted for the emission year 2015 for CO₂, CH₄ and N₂O. Sensitivity to methodological choices is also discussed. The methodology presented here has already applied to EDGAR and discussed in the scientific literature in comparison to other methods (Choulga et al., 2020), to other inventories (Petrescu et al., 2020), to assess the uncertainty of the EDGAR-FOOD inventory (Crippa et al., 2021), applied to specific sectors (Muntean et al., 2021), trend analysis of global GHG emissions and to communicate with the policy makers and the public (Crippa et al., 2019, 2020c).

2 METHODOLOGY

EDGAR is a ‘bottom-up’ model for estimating emissions, relying on a large spectrum of AD covering human activities with a high degree of detail. AD are combined with EFs to yield the emission, per source, and country. For example, for combustion sources AD consist of fossil fuel consumption while the EF is the amount of emission produced per unit of activity. In this case the emission is typically obtained simply by multiplying AD by EF, while other sources (e.g. waste) require more sophisticated models.

AD are primarily retrieved from international statistics, complemented, when necessary, with information (e.g. trends) from other sources, such as scientific literature and national data. The quality, consistency, and comparability of AD through time and space are the essential features defining the quality of an emission database.
Default EFs compiled by IPCC Guidelines *(IPCC Guidelines, 2006, hereafter referred to as IPCC-06)* are adopted by EDGAR for most sources and countries, supplemented by information from scientific literature, and other references for specific processes and/or countries. *Janssens-Maenhout et al. (2019)* produced a detailed description of data providers and methodological choices for the GHGs emissions of EDGAR. Further information on methodological aspects of data collection and sources are given by Crippa et al. (2020a).

This study addresses the uncertainty of the anthropogenic sources covered by EDGAR, which might be not exhaustive. Therefore, nothing can be said about the uncertainty stemming from source categories not currently encompassed within the inventory (e.g., fugitive CO₂ from low temperature oxidation of coal mines, fugitive CH₄ from managed wetlands, N₂O from crab ponds as part of aquaculture). Uncertainty assessment of spatially distributed sources (emission gridmaps) is not within the scope of this study.

#### 2.1 EMISSIONS AND THEIR UNCERTAINTIES

The uncertainty of AD (uₐₐ) collected by international agencies or organisations (e.g., the Food and Agriculture Organization (FAO), International Energy Agency (IEA)) is of statistical nature, stemming from incompleteness, representativeness of sampling, imputation of missing data, extrapolation (e.g., projecting to future years) *(Rypdal and Wininwarter, 2001; Olivier, 2002; IPCC-06)*. Other aspects to take into consideration when compiling a global inventory are the degree of wealth of a country as well as the year under study. Less developed countries and countries whose economy has fully developed in recent years, are more probable to have not yet developed a reliable statistical system. Similarly, AD of countries with transitional economies are expected to be more accurate for recent years *(Janssens-Maenhout et al., 2019)*.

Uncertainty in EF (uₑₑ) has many sources, as for instance: degree of representativeness of the limited number of observations underlying the EF; for the activity that is addressed, including under-representation of operating conditions; inaccuracy of assumptions and/or of source aggregation (e.g., assumption of constancy in time); bias, variability and/or random errors. Due to the non-statistical nature of uₑₑ, its quantification eludes a general methodological approach. IPCC adopts a tiered approach for estimating uncertainty, accounting for different levels of sophistication (IPCC-06). Tier 1 uncertainties on default EFs are based on expert judgement, which often offers a range of uncertainties for a given process, source, and/or fuel. Higher tiers (up to Tier 3) offer more elaborate estimates, based on localized measurements/ad-hoc experiments on specific emission factors and for specific processes.

Further, the model used to build emission inventories based on activity statistics may be too simplified (e.g., based on linearization and/or linear regression due for example to poor understanding, lack of data, etc.), and may not fully capture the complexity of a given emission process. These ‘model’ errors are difficult to be assessed in isolation from other sources of uncertainty, and are generally attributed to uncertainties in EFs *(Rypdal and Wininwarter 2001; Cullen and Frey, 1999)*.

This study reflects the methodological approach of EDGAR adopting default EFs, thus associated with Tier 1 uncertainty estimates. The term ‘uncertainty’, in this study as in similar ones *(Rypdal and Wininwarter, 2001; Olivier, 2002; Janssens-Maenhout et al., 2019)*, is used in a rather broad sense, lumping together all mentioned sources of errors due to current limited knowledge to distinguish among them. After IPCC introduced quantitative uncertainty in GHG inventories, the inventory uncertainty is usually expressed as two standard deviations, approximately corresponding to 95% confidence for a variable with a normal distribution (i.e., the uncertainty reflects the square root of the variance of the variable, multiplied by a coverage factor of 2 to provide a confidence interval of 95%).
Finally, the uncertainty tackled here shall not be confused with the variability stemming from a range (or ensemble) of estimates. The variability is used as proxy of structural uncertainty in the faith that a range of models using diverse underlying assumptions would span the true uncertainty space. However, the estimates are seldom ‘diverse’ as they build up from same data/assumptions (sometimes different versions of the same model are used) leading to overconfident estimates (Solazzo et al., 2018).

### 2.1.1 UNCERTAINTY IN ACTIVITY DATA

Table 1 summarizes the uncertainty for AD. When two values are listed (e.g. ±5%; ±10%), the lower uncertainty value (i.e. ±5%) is assigned to countries with developed economy, while the larger values (i.e. ±10%) to countries with less developed economy or with economy in transition.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>According to IPCC-06, $u_{AD}$ for fuel combustion activities (mostly derived from IEA statistics) are estimated with high confidence (5 to 10% uncertainty). The same uncertainty range is estimated for fugitive emissions (referring to venting and flaring during oil and gas production). $u_{AD}$ in the residential (10 to 20%) and in the aviation and navigation (5 to 25%) sectors are assumed more conservative to account for the under-representativeness of the sample and for the difficulty of distinguishing between domestic and international fuel consumption (IPCC-06). For combustion processes using biofuels, the statistics is less robust. Olivier (2002) suggests $u_{AD}$ of 30% for industrialised countries and 80% for less developed ones (based on IPCC-06 recommendations). Recent updates (Andreae, 2019) confirm these estimates.</td>
</tr>
</tbody>
</table>

Uncertainty for some chemistry production processes and waste is calculated on the total emission rather than on AD and EF separately, and is discussed later. The waste sector also utilizes a slightly more elaborated emission estimate model than the simple multiplication of AD and EF. It assumes that emissions are not instantly released into the atmosphere, but are accumulated and continue to emit even several years after their disposal. The model for the waste sector depends on several parameters and assumptions (detailed in section 3.1.5).

### 2.1.2 UNCERTAINTY ON EMISSIONS FACTORS

Tables 2 and 3 define the uncertainties of EFs for CO₂, and for CH₄ and N₂O, respectively. Uncertainty of EFs for CO₂ is determined by the carbon content of the fuel and is relatively smaller and determined with higher level of accuracy than uncertainty of EFs for CH₄ and N₂O. Moreover, use for CH₄ and N₂O lumps several sources of uncertainties, as mentioned earlier.

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>As adverted before, $u_{EF}$ are founded on Tier 1 estimates by IPCC-06, which are based on expert judgments and, as such, they vary over wide ranges to account for a variety of conditions. For instance, $u_{BF}$ for N₂O (agriculture and energy sources in particular) clearly reflect the large temporal variability and spatial heterogeneity of these processes.</td>
</tr>
</tbody>
</table>

### 2.2 EMISSION AGGREGATION AND UNCERTAINTY PROPAGATION

The vast majority of EFs in EDGAR are based on IPCC Tier 1 estimates (especially for combustion sources) to ensure completeness accomplished through the inclusion of all relevant sources for a given year.
The adoption of comparable methods for source emissions and consistency implies that the uncertainties of the final emission estimates are interlinked, as they stem from the same methodology. When emissions are combined/aggregated, this lack of independence cannot be neglected, and the following assumptions are made:

- emissions uncertainty ($u_{\text{EMI}}$) is the sum of the squares of the uncertainty of AD ($u_{\text{AD}}$) and the uncertainty of EF ($u_{\text{EF}}$) (see Eq. 1);
- uncertainties of different source categories are uncorrelated (e.g. waste and agriculture);
- subsectors of a given emission category for CH$_4$ and N$_2$O are fully correlated, thus the uncertainty of the sum is the sum of the uncertainties;

- when dealing with CO$_2$, full correlation is assumed for energy combustion sources sharing the same emission factor (fuel-dependent);
- aggregated emissions from same categories but different countries are assumed to be fully correlated, unless the emission factor is country-specific, or derived from higher tiers (i.e. emissions are not derived from default EF defined by IPCC but are retrieved by other sources and are specific to that country/process);
- when uncertainty is provided as a range (e.g. for the energy sector, IPCC-06 recommend that the CH$_4$ EFs are treated with an uncertainty ranging from 50% to 150%), the upper bound of the range is assigned to countries with less developed statistical infrastructure and the lower one to countries with more robust statistical infrastructure.

Conditions a) and b) match the suggestion of the uncertainty chapter of the IPCC guidelines (IPCC-06, Chapter 3), whilst the latter two conditions are more cautious formulations of the error propagation to account for covariances. More explicitly the uncertainty of the emission, $u_{\text{EMI}}$, due to multiplying AD by EF is calculated as:

$$u_{\text{EMI}} = \sqrt{(u_{\text{EF}}^2 + u_{\text{AD}}^2)} \quad \text{EQ. 1}$$

The uncertainty on the emission, $u_{\text{EMI}}$, due to adding emissions is calculated as:

$$u_{\text{EMI}} = \sqrt{\sum (E_{\text{MI}}, i \cdot u_{\text{EMI}}, i)^2 / \sum |E_{\text{MI}}, i|} \quad \text{EQ. 2}$$

That is, basically, the squared sum of the uncertainty of each emission process normalised by the sum of emissions, which assumes that all emission sources are uncorrelated (IPCC-06). However, in general, the variance of the sum of any two terms $x_1$ and $x_2$ having variances of $\sigma_1$ and $\sigma_2$ is $\sigma_{\text{sum}}^2 = \sigma_1^2 + \sigma_2^2 + 2\text{cov}(x_1,x_2)$. Since the covariance can be expressed as $2\text{cov}(x_1,x_2) = 2r\sigma_1\sigma_2$, where $r$ is the coefficient of correlation, when $r = 1$ (full correlation), the variance of the sum becomes the linear sum of the two variances:

$$\sigma_{\text{sum}} = \sigma_1 + \sigma_2 \quad \text{correlated } r=1 \gtrless \sqrt{\sigma_1^2 + \sigma_2^2} \quad \text{uncorrelated } r=0 \quad \text{EQ. 3}$$
Therefore, for fully correlated variables, the uncertainty of their sum is simply the sum of their uncertainties.

When uncertainties are larger than 100%, Eq. 2 tends to underestimate the uncertainty and a correction factor $F_C$ is recommended (IPCC-06), so that the uncertainty on the emission is:

$$ F_C = \frac{u_{EMI,C}}{u_{EMI}} $$

where $u_{EMI,C}$ is the correction to be applied to the uncertainty estimated from error propagation. Eq. 4 is used for multiplicative or quotient terms in the range $u_{EMI,C} \in [100\%, 230\%]$ (Equation 3.3, IPCC-06 Volume 1 Chapter 3). The effect of $F_C$ is to return larger uncertainties (see e.g. Choulga et al., 2020). The use of $F_C$ is based on the work by Frey (2003) to account for the error introduced in the approximation of the analytical method compared to a fully numerical one (based on Monte Carlo analysis). The error in the approximation increases with the uncertainty, and thus the correction factor $F_C$ is needed when dealing with large uncertainties (Frey, 2003). The analysis presented in this study takes into account for the correction factor $F_C$ (unless specifically indicated) and for simplicity the ‘C’ is dropped in $u_{EMI,C}$ to yield $u_{EMI}$.

This study assumes that uncertainties are normally distributed, unless specifically indicated by IPCC-06. The distribution is transformed to log-normal after the aggregation to avoid that the emissions take negative, unphysical values when uncertainty is large. Hence, the probability distribution function (PDF) is transformed to lognormal with the upper and lower uncertainty range defined according to IPCC-06:

$$ u_{EMI} = \frac{1}{EMI} \left( \exp(\mu_g) \pm 1.96 \ln(\sigma_g) \right) - 1 $$

where $\mu_g$ and $\sigma_g$ are the geometric mean and geometric standard deviation about $EMI$, the mean emission.

According to IPCC-06, the contribution to variance, $var\ share_g$, of a specific emission process $s$ emitting $EMI$, to the uncertainty of the total emissions $EMI_{tot}$ is calculated as:

$$ var\ share_g = \frac{u_{EMI,s}^2 + EMI_s^2}{EMI_{tot}^2} $$

### 3.2.1 ADDITIONAL REMARKS

The assumption of correlation between subcategories (or fuel for energy sector emitting CO$_2$) and between countries for the same category (or fuel for energy-CO$_2$) is introduced to ensure that the uncertainty of sources sharing the same methodology for estimating the EF is propagated in case of aggregation. If the same methodology is applied to estimate the emission for a given category and for a group of countries, then the correlation is kept when calculating the total emission of that group of countries for that category. Similar assumptions were adopted by e.g., Bond et al. (2004) and Bergamaschi et al. (2015) (though for different inventories). This is a direct implication of the consistency and cross-country comparability of EDGAR, that adopts Tier 1 EFs defined by IPCC-06.
for most of the inventory. By contrast, if each country follows diverse methods to estimate the EFs for a given source category, $\sigma_E$ stemming from that methodology does not co-vary when calculating the total of that category, and thus Eq. 2 holds. Some further considerations:

- the assumption of source/country correlation is the main difference between the uncertainty estimated in this study and the uncertainty reported by, e.g., Petrescu et al. (2020) for EU27+UK, where no correlation was assumed, although not all countries developed independent methods to estimate EFs.
- the choice of assuming ‘full’ correlation (i.e., correlation coefficient of one) is conservative in the sense that it returns the upper bound of $\sigma_E$, and is motivated by two main reasons: it simplifies the calculation (see Eq. 3), and there are no indications now to better estimate $\rho$.
- EDGAR does include country-specific EFs for some processes and countries. Those are retrieved from the scientific literature or derived from technical collaborations, and through the continuous updates over the last two decades (e.g., EFs for cement production are computed including information on country-specific clinker fractions. EFs for landfills consider the country-specific waste composition and recovery. EFs for enteric fermentation of cattle include country/region-specific information on milk yield, carcass weight and many other parameters, etc.). These instances are flagged in our methodology, and the $u_E$ is not propagated when aggregating these sources.

3. Uncertainty in Emission Sectors

3.1 Emissions from CO$_2$, CH$_4$, and N$_2$O

3.1.1 Power Industry Sector

IPCC sector 1.A includes the EDGAR categories related to combustion of fossil and biofuels for energy production (ENE), manufacturing (IND), energy for buildings (RCO), oil refineries and transformation industry (REF, TRF), aviation (TRN aviation), shipping (TRN ship), and road transport (TRO).

Emissions from biofuel burning (e.g., wood) in sector 1.A are considered carbon neutral and are calculated for CH$_4$ and N$_2$O only.

EDGAR adopts AD statistics of fossil fuel combustion compiled by the IEA (IEA, 2017) for developed and developing countries, integrated with data from EIA (2018) for biofuels.

TABLE 4.

The share of GHGs emissions from industrialised and developing countries is reported in Table 4 to aid later interpretation of the uncertainty shares. In fact, in countries with developed economy (Table S1) energy statistics are considered to have lower uncertainty than in countries in development (Olivier 2002). IPCC suggests $u_E$ for the power industry ranging between 5 to 10%. We have assigned 5% to industrialised countries and 10% uncertainty to developing countries to account for less robust census capability. IPCC-06 provides fuel-dependent $u_E$ for CO$_2$ (Table 2), which have been mapped to match the fuels in each EDGAR emission category. $u_E$ for CO$_2$ are relatively small as reflected by the (well known) carbon content of the fuel.

For CH$_4$ and N$_2$O, EFs are more uncertain than for CO$_2$. IPCC-06 suggests a wide range of $u_E$ for the whole energy sector, ranging between 50% and 150% for CH$_4$ and between one tenth and ten times the mean emission value for N$_2$O. These estimates are provided by expert judgement based on the reliability of current estimates. The reasons for such high uncertainty are those mentioned before, i.e., lack of understanding of emission processes and of relevant measurements, uncertainty in measurements, poor
representativeness of the full range of operating conditions. EFs for biofuels combustion are highly uncertain, estimated in the range 30% (Andreae and Merlet, 2001) to 80% (Oliver, 2002). Recently, Andreae (2019) has reviewed $u_{EF}$ to less than 20% (6-18% for CH$_4$ from the major burning categories savanna, forests, and biofuel). The uncertainty of processes using biofuels is calculated separately and then combined with the fossil fuel uncertainty, assuming no correlation, see Eq. 2).

FIGURE 1.

Emissions of CO$_2$ account for over 90% of world’s total GHG emissions from fuel combustion, and are assessed with high degree of confidence (Figure 1a,b,c) due to the accuracy of $u_{EF}$ reflecting the carbon content of the fuel. Thus, the share of emission for each subcategory (manufacturing, transformation and power industry, oil refinery, residential heating, road and non-road transport) is mirrored by the share each category contributes to the sector uncertainty (Figure 2), although with some notable exceptions for non-road transport in Brazil (large share of highly uncertain domestic aviation and inland water shipping), and transformation industry in Russia (share of emission and uncertainty of ~10% and ~37%, respectively).

FIGURE 2.

The very low confidence in N$_2$O emissions is responsible for almost 50% of world’s total uncertainty (Figure 1f) although N$_2$O only accounts for a minor portion of total emissions in this sector (less than 1%). According to, e.g., Lee et al. (2013), the suggested IPCC-06 uncertainty on power plant emission of N$_2$O might be too high (the authors report a range of −11.43% and ±12.86% for combined-cycle power plant in Korea). An alternative $u_{EF}$ estimation for N$_2$O in the fossil fuel combustion sector is set in the range ±50% (developed countries) to ±150% (countries with economy in development). This choice also reflects previous uncertainty estimation by Olivier (2002).

The N$_2$O emission uncertainty and the N$_2$O contribution to uncertainty in sector 1.A become as shown in Figure 3.

FIGURE 3.

The uncertainty distribution (Figure 3) and relative contribution reflect the weight of the component GHGs and the world’s total uncertainty (10%) is only slightly larger than the uncertainty of CO$_2$ (7%, Figure 1a,b,c). Adopting the $u_{EF}$ of 50-150% for N$_2$O in sector 1.A reflects the large uncertainty associated with this sector and allow comparability/aggregation with other gases (Figure 3b).

3.1.2 FUGITIVE EMISSIONS FROM COAL, OIL AND NATURAL GAS

Fugitive emissions from solid fuels (mainly coal, 1.B.1) and from oil and natural gas (1.B.2) are covered by the EDGAR’s categories REF, TRF and by fuel exploitation PRO. As pointed out in IPCC-06, uncertainty in the fugitive emissions sector arises from applying the same EF to all countries (Tier 1 approach) and from uncertainty in the emission factors themselves.

AD for coal statistics is a collection of products (full details are provided by Janssens-Maenhout et al. (2019) and references therein): the World Coal Association (2016); JEA (2017) for exploration of gas and oil; UNFCCC (2018) and CIA (2016) for transmission and distribution; JEA (2017) for venting and flaring, complemented with data from GGFR/NOAA data (2019) and Andres et al. (2014). According to Olivier (2002), $u_{EF}$ for sector 1.B lies within the range ±5 to ±10%, which is aligned with the estimates provided by IPCC-06.
Fugitive emissions from solid fuels (1.B.1) in EDGARv4 and v5 are dealt with by considering emission factors from IPCC-06, supplemented with EMEP/EEA (2013) Guidebook for coal and UNFCCC (2018). For oil and natural gas (1.B.2), we use information from the IPCC-06, supplemented with data of UNFCCC (2014). While gas transmission through large pipelines is characterised with relatively small country-specific emission factors of Lelieveld et al. (2005), much larger and material dependent leakage rates of IPCC-06 were assumed for gas distribution. For venting processes EFs for CH\textsubscript{4} are based on country-specific UNFCCC (2014) data for reporting countries (and the average value as default for all other countries) (Janssens-Maenhout et al., 2019).

IPCC-06 provides a detailed synthesis of uncertainty associated with EFs for sectors 1.B.1 and 1.B.2, distinguishing between developing and developed countries (Tables 4.2.4 and 4.2.5 of IPCC-06, chapter 4). u\textsubscript{EF} is the same for CO\textsubscript{2} and CH\textsubscript{4}, while is larger for N\textsubscript{2}O. A summary of uncertainty ranges is provided in Table 3.

Uncertainties in the 1.B.1 sector depend on the type of mining activity: ‘surface’ (surf), ‘underground’ (und) and ‘abandoned’ (abandon). u\textsubscript{EF} for these sectors can be rather large (>100%), as detailed in Table 3, according to IPCC-06 and in line with Olivier (2002). For 1.B.2, the distinction is made between leakage in production (prod), transmission and distribution (trans), and venting/flaring (vent). The uncertainty is estimated as large as three times the average emission value for some instances (Table 3) for CH\textsubscript{4} and CO\textsubscript{2}, and up to 100% for flaring N\textsubscript{2}O emission. We note that while some AD are known or retrievable through various governmental agencies (e.g. number of gas production wells, miles of pipelines, number of gas processing plants), other activity data (e.g., storage tank throughput, number of various types of pneumatic controllers, and reciprocating engines) are more uncertain. As reported by EPA, ‘petroleum and gas infrastructure consist of distinct emission sources, making measurement of emissions from every source and component practically unfeasible’ (EPA, 2017).

**FIGURE 4.**

The fugitive emission sector is dominated by CH\textsubscript{4} emissions and this is reflected in the contribution to the total uncertainty of GHG emission from sector 1.B (Figure 4c). The upper world uncertainty estimate exceeds 110%, almost entirely due to CH\textsubscript{4} emissions. For the USA, upper uncertainty estimates for oil and natural gas (Figure 4c) of 23% is slightly less than the EPA’s upper estimate of 30% for the natural gas system (EPA, 2017) and that of Littlefield et al. (2017) of 29%, while for the petroleum system the EPA’s uncertainty is much larger (149%), possibly due to higher u\textsubscript{AD}.

The uncertainty of individual countries mirrors the distinction made between developed and developing countries, mostly visible for fugitive emissions from oil and natural gas (Figure 4c) but also in the detailed use provided by IPCC-06 for the various emitting stages of extraction, distribution, transport, and storage. The composition of emissions for the five top emitters in sector 1.B.2.b can be used to illustrate this aspect.

**TABLE 5.**

The USA and Russia have country-specific EFs, which are defined for all stages of the fugitive emissions from natural gas, and therefore the accuracy is high. Iran, Saudi Arabia, China have a very large share of emissions due to the production stage of natural gas (approximately 85%, 97%, 76%, respectively, Table 5), to which u\textsubscript{EF} = ±75% applies, and a much lower share of emissions apportioned to the other stages (i.e. transmission and distribution), approximately 10% due to gas distribution with an uncertainty of ±40% to +500% (including the correction factor Eq. 4), contributing to the very low confidence in the emission estimate shown in Figure 4c, compared with the medium confidence for...
USA and Russia, to which country-specific EFs are applied (±25%) (Table 3). The high uncertainty in the transmission/distribution sectors is the main cause for the difference in uncertainty apportionment.

Variability of bottom-up estimates of CH4 emissions from coal mining (-29%, +43%) and natural gas and oil systems (-16%, +15%), as recently reported by Saunois et al. (2020), stems from methodologies and parameters used, including emission factors, "which are country- or even site-specific, and the few field measurements available often combine oil and gas activities and remain largely unknown" (Saunois et al., 2020). The authors reported examples of very large variability of EFs between inventories, even of 2 orders of magnitude for oil production and by one order of magnitude for gas production. Moreover, large uncertainties in emissions of CH4 from venting and flaring at oil and gas extraction facilities were reported by e.g. Peischl et al. (2015). Gas distribution stage is a further large source of uncertainty, in particular in countries with old gas distribution city networks using steel pipes now distributing dry rather than wet gas, with potentially more leakages (Janssens-Maenhout et al., 2019). Analysis based on inversion modelling by Turner et al (2015) found, for the North America region an error variability of -43% to 106% (with respect to the prior estimate based on EDGAR v4.2) attributed to emissions from oil and gas. Hence, the uncertainty in Figure 4c might be too low for industrialised countries. For completeness, we show an alternative application of uncertainty ranges for sector 1.B.2 (oil and gas), as suggest by Olivier (2002), assigning uAD = ±5% and ±15% (industrialised and developing countries, respectively) and uEF = ±100% to all countries and uEF of 50% to countries for which EF are specifically estimated (Tier 3).

**FIGURE 5.**

The resulting distribution (Figure 5) reflects the comparable uncertainty of these emissions across countries. Global uAD is of approximately 100%, thus slightly less than the uncertainty obtained by applying the IPCC-06 recommendations (122%, Figure 4c).

3.1.3 INDUSTRIAL PROCESSES AND PRODUCT USE (IPPU)

IPPC category 2 covers non-combustion emissions from industrial production of cement, iron and steel, lime, soda ash, carbides, ammonia, methanol, ethylene, adipic and nitric acid and other chemicals and the non-energy use of lubricants and waxes (Janssens-Maenhout et al., 2019). The EDGAR sectors CHE (production of chemicals), FOO (food production), PAP (paper and pulp production), IRO (iron and steel), non-energy use of fuels (NEU), non-ferrous metal production (NFE) and non-metallic minerals production (NMM) cover the industrial process emissions.

Activity statistics for industrial processes are retrieved from several reporting providers, as detailed by Janssens-Maenhout et al., 2019 and Crippa et al., 2019. For this class of processes uAD are higher than uEF due to the deficiency or incompleteness of country specific data and reluctance by companies to disclose production data. CO2 emissions in EDGAR are based on Tier 1 EF for clinker production, whereas cement clinker production is calculated from cement production reported by USGS (2014). The fraction of clinker is based on data reported to UNFCCC for European countries, to the China Cement Research Institute (www.ccement.com; yjy.ccement.com/) and the National Bureau Statistics of China (for historic years) for China and to the ‘getting the numbers right’ for non-Annex I countries (https://gccassociation.org/gnr/). According to IPCC-06, the uncertainty for cement production stems prevalently from uAD, and to a lesser extent from uEF for clinker (IPCC-06, chapter 2). For Tier 1, the major uncertainty component is the clinker fraction of the cement(s) produced and uAD can be as high as 35%. We assume uAD of 11% to 60% depending on the accuracy of clinker data.

As for cement, the uAD for lime outweighs uAD due to lack of country specific data. We assume uAD of ±35% and uEF = ±3%. For glass, glass production data are typically measured accurately as reflected by
Production of ammonia, nitric and adipic acid as well as caprolactam, glyoxylic and glyoxylic acid is known with high degree of accuracy and $u_{AD}$ for these processes can be estimated as $\pm 2\%$. The corresponding $u_{EF}$ is reported in Table 2, Table 3 and is derived from expert judgment elicitation and reported in IPCC-06 ($u_{EF}^{\text{Ammonia}} = \pm 7\%$; $u_{EF}^{\text{Nitric Acid}} = \pm 20\%$; $u_{EF}^{\text{Carbon}} = \pm 10\%$). For petrochemical and carbon black production (methanol, ethylene, ethylene dichloride, vinyl, acrylonitrile, carbon black), IPCC-06 provides reference values for $u_{AD}$ associated to these processes ($IPCC-06$, Volume 3, Chapter 3, Table 3.27), based on expert judgments. The values are reported in Table 3, ranging from $\pm 10\%$ for CH$_4$ emission for ethylene production to $\pm 85\%$ for CH$_4$ emission from carbon black production.

As summarised in Table 1, the AD for iron and steel (including furnace technologies) production are considered very accurate, with $u_{AD} = \pm 10\%$, and for ferroalloys $u_{AD}$ is set to $\pm 10\%$ for industrialised countries and $u_{AD} = 20\%$ for developing countries, based on own judgment (IPCC-06 suggests $u_{AD} = \pm 5\%$). The data for iron production are updated monthly using data from the World Steel Association (WSA, 2019), while for ferroalloys data are extrapolated using trends from USGS commodity statistics (USGS, 2016). $u_{EF}$ is equal to $\pm 25\%$.

Production data for aluminium, magnesium, zinc, and lead are deemed accurate within 2% to 10% (Table 1). For aluminium, the reactions leading to CO$_2$ emissions are well understood and the emissions are very directly connected to the quantity of aluminium produced (IPCC-06), and $u_{EF}$ is assumed within 10%. The $u_{EF}$ associated with CO$_2$ emitted from magnesium production is also well understood and is assumed within 5%. Lead and zinc production have higher $u_{EF}$ (50%) associated with default emission factors (Tier 1), and of 15% if country specific data are adopted (Tier 2). CO$_2$ emissions for non-energy use of lubricants/waxes (like petroleum jelly, paraffin waxes and other waxes, classified under IPCC sector 2.D.2 and corresponding to EDGAR sector NEU) are assumed highly uncertain ($u_{EF}$ of 100%; $u_{AD}$ of 5% to 15%) due to the lack of accurate information and to country specific operating conditions.

**FIGURE 6.**

CO$_2$ emissions in sector 2 are one and two orders of magnitude higher than N$_2$O and CH$_4$ emissions respectively (Figure 6). Nearly 50% of CO$_2$ emissions in this sector originate from cement production. The accuracy ranges from medium-high to high for all top emitters, and the global uncertainty is of 12%. For N$_2$O, the main source (85%) is the production of nitric and adipic acid, which results in medium-high accuracy both country wise and globally. Finally, emission of CH$_4$ is more uncertain due to the large $u_{EF}$ of carbon black and methanol production, which account for $\sim 52\%$ of global CH$_4$ emissions in the IPPU sector.

### 3.1.4 AGRICULTURE

Agriculture related activities in EDGAR cover partially the IPCC category 3 (agriculture, forestry and land use), including enteric fermentation (ENF, corresponding to 3.A.1), manure management (MNM, 3.A.2), waste burning of agricultural residues (AWB, CRP, corresponding to 3.C.1.b – biomass burning of cropland), direct N$_2$O emissions from soil due to natural and synthetic fertiliser use (corresponding to 3.C.4), indirect N$_2$O emissions from manure and soils (corresponding to 3.C.5 and 3.C.6), urea and agricultural lime (AGS.LMN and AGS.URE, corresponding to IPCC codes 3.C.2 and 3.C.3), and rice
For sectors ENF and MNM, EDGAR follows IPCC-06 for estimating emissions, with animal counting data from FAO (2018). For ENF, uncertainty in AD is due to cattle numbers, feed intake, and feed composition, while for MNM the distribution of manure (volatile solids) in different manure management systems is also a source of uncertainty. \( u_{AD} \) for these sectors is estimated of \( \pm20\% \) to account for uncertainty of the manure management system usage, of lack of detailed characteristics of livestock industry, information on how manure management is collected, and lack of homogeneity in the animal counting systems (IPCC-06; Olivier, 2002). The estimate is slightly higher than \( u_{AD} \) from other USA studies for ENF (EPA, 2017; Hristov et al., 2017), whilst for MNM \( u_{AD} \) of \( \pm20\% \) might be underestimated according to, e.g., Hristov et al. (2017). EFs are calculated following IPCC-06 methodology, using country specific data of milk yield and carcass weight integrated with trends from FAO (2018) for cattle, and using regional EFs for livestock. Tier 1 \( u_{EF} \) for ENF and MNM is estimated to be larger than \( \pm50\% \) (with a minimum of 30\%) unless livestock characterisation is known with great accuracy, in which case Tier 2 uncertainty can be \( \pm20\% \) (IPCC-06).

AD for burning of agriculture waste (AWB.CRP) can be highly uncertain, especially in developing countries, due to several factors including the estimates of the area planted under each crop type for which residues are normally burnt and the fraction of the agricultural residue that is burnt in the field. EDGAR estimates the fraction of crop residues removed and/or burned using data from Yevich and Logan (2003) and from official country reporting. Uncertainty is deemed very high, in the range \( u_{AD} = 50 \text{ to } 100\% \) (Olivier, 2002; Olivier et al., 1999a). EFs for this sector are obtained from the mass of fuel combusted, provided by IPCC-06 as default (Tier 1) EFs for stationary combustion in the agricultural categories, and are estimated with an uncertainty of \( \pm60\% \) to \( +275\% \) for \( N_{O} \), and \( \pm50\% \) to \( \pm150\% \) for \( CH_{4} \) according to the uncertainty for combustion processes.

Emissions from rice cultivation are relevant to \( CH_{4} \). According to the last release of EDGAR, in 2015 almost 10\% of total \( CH_{4} \) emissions were due to rice cultivation. Default baseline EF for rice cultivation has an uncertainty in the range \( -40\% \) to \( +70\% \), which has been substantially reviewed in the IPCC refinement (2019), both in terms of EF value and of uncertainty. The refinement also gives regional-dependent EF and uncertainty ranges, but those have not been implanted yet in EDGAR, therefore we refer to the IPCC-06 guidelines. In EDGAR the baseline EF is multiplied by a set of scaling factors that account for the water regimes before and during the cultivation period: upland (UPL, never irrigated), irrigated (IRR), rain fed (RFN) and deep water (DWP), which are assigned the following uncertainty (derived from IPCC-06): IRR = \( -20\% \) to \( +26\% \); UPL = \( 0\% \); RFN and DWP = \( -22\% \) to \( +26\% \). Organic amendments and soil type are not included. The AD consist of cultivation period and annual harvested area for each water regime and are derived from FAO (2011) and are complemented with data from IRRI (2007) and IIASA (2007). We assume \( u_{AD} \) of 5\% to 10\% (Olivier, 2002). All the conditions together yield an uncertainty range of \( -0.45\% \) to \( +75\% \) for RFN, DWP and IRR, and of \( -0.41\% \) to \( +70\% \) for UPL.

**Figure 7.**

AD for sectors 3.C.2 (CO2 emissions from liming), 3.C.3 (CO2 emissions from urea application), are derived from FAO (2016), and from official country reporting. Uncertainty of emissions of CO2 from lime (urea) fertilization stems from uncertainties in the amount of urea applied to soils and from the uncertainties in the quantity of carbonate applications that is emitted as CO2. \( u_{AD} \) is assumed of 20\% (Olivier et al., 1999a) to account for uncertainty in sales, import, export and usage data adopted to derive...
the AD. EFs are derived from IPCC-06 Tier 1, assuming that all C in urea is lost as CO2 in the atmosphere, which might give rise to systematic bias. uEF is assumed between ±50% and ±100%.

Sectors 3.C.4, 3.C.5, 3.C.6 cover direct and indirect N2O emissions from managed soils and manure management. AD are taken from FAOSTAT (2016) and UNFCCC (2018). Nitrogen from livestock data for developed countries is derived from the CAPRI model (Leip et al., 2011) and can be considered as Tier 3 level accuracy. Indirect N2O emissions are due to leaching and runoff of nitrate and are subject to various sources of uncertainty (both AD and EFs) due to natural variability and to the volatilization and leaching factors, poor measurement coverage and under-sampling as well as due to incomplete/ inaccurate/ missing information on observance of laws and regulations related to handling and application of fertiliser and manure, and changing management practices in farming (IPCC-06). For these sectors, uAD is estimated ±20% and uEF in the range ±65% to ±200% according to IPCC-06).

Studies by e.g. Philibert et al. (2012) and Berdaniere and Coron (2012) suggest that the uncertainty of N2O emissions due to N fertilization can be as lower as up to a factor 5.

The large variation of N2O emissions in time and space is well recognised (e.g. Stenhf and Bonneman, 2006). Spatial heterogeneity, in particular, is largely driven by soil properties, and the influence of soil properties changes with scale and is responsible for the large confidence intervals given for the IPCC EFs (Milne et al., 2014).

With a few exceptions, the confidence in emission estimates from agriculture varies between medium and low for CO2 and CH4 (Figure 7a,b) depending on the composition of the agricultural sources and on the accuracy assigned to the specific country (developing vs industrialised). N2O (Figure 7c) emissions are very uncertain (in excess of 300%), which is reflected in the global share of uncertainty (over 90%, though the share of global N2O emissions does not exceed 30%, Figure 7d).

For the UK, Milne et al. (2014) estimated a 95% confidence interval of −56% to +139%, Brown et al. (2012) of −93% to +2533%, whereas Monni et al. (2007) of −52% to +70% for Finland (but based on older and more conservative IPCC guidelines). Our uncertainty estimates for the UK for sectors 3.C.4, 3.C.5, 3.C.6 combined is of −74% to 305% (as direct effect of assuming full correlation, in fact if the three sectors were considered to be uncorrelated, the 95% confidence interval for the UK would be −59% to +259%, which is in line with the other estimates).

Figure 8.

Uncertainties due to rice cultivation and enteric fermentation outweigh the uncertainty from other sources, being the dominant emission shares over the emissions from burning of crop residues (which has higher uncertainty but low impact on overall emission) (Figure 8). Agricultural uncertainties in China are attributable to rice cultivation for ~30%, whilst rice emission accounts for less than 60% of the agriculture total. Similarly, the uncertainty due to enteric fermentation dominates the USA agriculture uncertainty (75% share).

3.1.5 Waste

The waste-related emissions in EDGAR correspond to IPCC category 4 (waste), including emissions from managed and non-managed landfills (SWD: solid waste disposal on land and incineration, categories 4.A, 4.B and 4.C), wastewater handling (domestic WWT.DOM and industrial WWT.IND, categories 4.D.1 and 4.D.2, emitting CH4 and N2O), and waste incineration (emitting CH4, N2O, and also CO2). Globally, the waste sector accounts for 4.4% of total GHG anthropogenic emission in 2015 and 21.5% of total anthropogenic CH4 emissions (Crippa et al., 2019).
In EDGAR, emissions are based on a combination of population and solid and liquid waste product statistic. CH₄ emissions from landfills are calculated following the first order decay model proposed by IPCC-06, which assumes that emissions do not occur instantaneously but are spread over several years.

The model depends on several parameters (Table 1 and Table 3), and the main factor in determining the CH₄ generation potential is the amount of degradable organic carbon (DOC) (IPCC-06; Olivier, 2002; Janssens-Maenhout et al., 2019). The average weight fraction of DOC under aerobic conditions is provided by the IPCC Waste Model for 19 regions, which has been used as the default for all countries. Moreover, the default parameters for the methane correction factor (MCF), constant (k) and the oxidation factor (OX) are adopted (full details in Table 1 of Janssens-Maenhout et al. (2019)). Each component of the waste model has been assigned a normal distribution using the 95% confidence interval defined in Table 1 and Table 3 and combined using a sample population of 10000 elements.

The range of overall uncertainty is between 35% and 134% for CH₄ and between 10% and 490% for N₂O.

For the incineration of waste, AD are derived from UNFCCC NIR, IPCC-06, country reports and scientific literature, extrapolated using population trends (e.g. for countries with scarce data on municipal solid waste), while for composting (category ‘other’), data are obtained from UNFCCC NIR for Annex I countries and scientific literature for developing countries and for India (Table 1 of Janssens-Maenhout et al., 2019 and references therein).

As detailed in Janssens-Maenhout et al. (2019), the IPCC-06 default values for wastewater generation and chemical oxygen demand (COD) are used to derive the total organically degradable material (TOW), differentiating by type of industry (meat, sugar, pulp, organic chemicals, ethyl alcohol). Population from UNHABITAT statistics (UNHABITAT, 2016) is used to derive country-specific percentages of population at mid-year residing in urban and rural areas, with low and high income, for calculating domestic wastewater. Different wastewater treatments are specified with technology-specific CH₄ emission factors. For domestic wastewater, the sewer to wastewater treatment plants (WWTP), sewer to raw discharge, bucket latrine, improved latrine, public or open pit and septic tank are distinguished. Uncertainty of domestic wastewater depends on the technology (sewer to raw discharge, bucket latrine, improved latrine) as specified in Table 1 and Table 3, and is composed of uncertainty in AD (population data, ±36%) and uncertainty on EF (±33% to 78%).

Uncertainty on AD for industrial wastewater data ranges between -56% to 103%, estimated using the IPCC-06 suggested values, which are in line with those provided by Olivier et al. (2002) (-50% to 100%). Uncertainty on EF includes 30% uncertainty for the maximum CH₄ producing capacity (parameter Bₐ) and uncertainty on the CH₄ correction fraction of ±50% to 100% (based on the range of default values for MCF provided by IPCC-06 in Table 6.8 of Volume 5).

Emissions of CH₄ from the waste sector is one order of magnitude higher than N₂O and two orders higher than CO₂ (Figure 9a,b,c) and although N₂O emissions are more uncertain, the share of uncertainty still reflects the share of emissions (Figure 9). The confidence in the emission estimates varies from medium to medium low for CO₂ (depending on the status of development of the country), from medium to very low for CH₄ (depending on the status of development of the country and on the composition of the waste sector, discussed below) and is very low for N₂O (due to high uₑₑ in waste water).

FIGURE 9.

The composition of the waste sector for CH₄ (Figure 10a) shows that there is a strong correspondence between the emissions share and the uncertainty share. For the USA, landfills emissions...
account for ~73% of waste emissions, and the uncertainty due to landfills is ~90%. In India, domestic wastewater accounts for over 85% of waste emissions, driving the overall uncertainty with 97%.

Worldwide, the CH₄ emission share from landfills and domestic wastewater is approximately equivalent (~44% and ~41%, respectively), whilst landfills have a relatively larger weight in the global uncertainty share (~55% and ~41%, respectively).

3.2 The Global and European Picture

The values in Table 6 summarise the global uncertainty ranges. First, the uncertainties are given for each sector and gas individually, then for the sum of the three GHGs for each sector, and finally for the sum of the three GHGs and for all the sectors together. The last row of the table, thus, is the overall EDGAR uncertainty on the worldwide GHG emissions.

TABLE 6.

Globally, while CO₂ is by far the largest emitted GHG (in excess of 75%) followed by CH₄ (19%), the main source of uncertainty (~50%) is N₂O (Figure 11a), followed by CH₄ (~29%). Agriculture alone accounts for 39% of the global uncertainty (Figure 11b) and is almost entirely due to N₂O as discussed earlier (Figure 8d) and energy accounts for 44% (almost half of the uncertainty for energy is due to N₂O, Figure 1f) and waste (11%, driven by CH₄ emissions, Figure 9d).

The picture is quite similar for EU27+UK (Figure 12) with the main difference being the larger uncertainty share of N₂O (~70%) due to the higher level of accuracy associated with CO₂ and CH₄.

4 Uncertainty Due to Methodology

The considerable number of ‘degrees of freedom’ influencing the uncertainty of an emission inventory such as EDGAR is itself a source of uncertainty originating from different methodological assumptions. As such, the structural uncertainty of emissions tackled in the previous section is subject to variability due to the sets of assumptions, methods, choices adopted for its quantification. It originates from lack of agreement/incomplete knowledge on the processes governing the emission sources and their representativeness. Such a methodological uncertainty reflects the judgment of the uncertainty emission compiler and can give rise to a significant share of the overall uncertainty estimate. For instance, two experts could suggest two different probabilistic models for the value of a certain emitting source, leading to a certain degree of variability in the PDFs of that source. Methodological uncertainty, thus, may arise from the assumptions adopted assessment, particularly when there are no clear guidelines or reference cases about methodological choices that allow comparability between evaluations.

One of the most impactful assumptions of this study is the correlation between subcategories/fuels and, for the same category/fuel, between countries. This has a profound impact on the uncertainty estimate, for example in inter-comparison studies where EDGAR’s uncertainties are shown next to other inventories whose uncertainty estimates do not account for correlation (e.g. Petrescu et al., 2020; Choulela et al., 2020).
The global weight of the correlation is reflected in the total of Figure 13, where the uncertainty ranges from 4% (no correlation) to above 20% for the correlated cases. The impact of assuming correlation of the uncertainties when aggregating the emissions of several countries outweighs any other assumptions. For instance, the assumption to constrain the N₂O uncertainty for energy in the range ±50% to ±150% has, globally, much lower impact over the total uncertainty (23% rather than 20%).

**Figure 14.**

As shown in Figure 14 for EU27+UK, the effect of correlation on the variability of the uncertainty is considerable. Emissions from the energy sector are estimated to be accurate, since the 95% confidence interval lies within 2% of mean value when no correlation is assumed across countries, and within 7% when the correlation is set to one. The uncertainty of 13% for the Tier1 ‘default case’ reflects the high share of uncertainty due to N₂O since the only difference between the ‘T1 default’ and ‘T1+OJ N₂O’ for energy is the upper limit of N₂O uncertainty to ±50% and ±150% (OJ: Own Judgment). The same argument applies to the other sectors, most notably to agriculture (130% vs 36%, with or without correlation), and is reflected in the total GHG emissions (15% vs 4%).

Important to notice that if EU27+UK report emissions as a single party, even Tier 1 propagation methods would return an accuracy comparable to the combination of independent estimates (i.e., as if all EU parties used independent, Tier 2 or 3 estimates of their emissions).

The comparison between the ‘default’ uncertainty ranges and ‘EDGAR in-house expert judgment’ for N₂O shows the impact of choices on the quantification of the uncertainty, contributing to enhance the uncertainty variability. The case of energy in Figure 14 is an example: the default uncertainty of 13% can vary as much as 46% (down to 7%) due to different judgments in estimating $\delta_{EF}$.

**5. Conclusions**

This study quantifies the structural uncertainty of the EDGAR inventory of GHGs. Given the widespread applications of EDGAR in many areas – modelling, policy, evaluation, planning – the qualification of its accuracy and quantification of its uncertainty are essential added values.

EDGAR is a consistent database based, predominantly, on Tier 1 methods to quantify emission from anthropogenic sources (on a three-level of sophistication, Tier 1 is the simplest). As such, the uncertainty analysis presented here follows the corresponding Tier 1 approach for uncertainties, also suggested by IPCC (2006; 2019) to assist in country reporting. Some additional assumptions have been put forward to allow for the simple Tier 1 uncertainty method to integrate with the EDGAR global database.

The global, comparable nature of EDGAR is one of its main attractivenesses. Zooming in individual countries, the accuracy of EDGAR cannot, in general terms, match that of the country’s inventory reporting panel who might adopt higher tiers for estimating emissions and uncertainties. Hence, it is when looking at cross-sector, gases and countries aggregation that the analysis presented in this study shows its benefits.

For the aggregation of emitting sources sharing the same underlying methodology, we have assumed that the uncertainty is amplified, and therefore the aggregation must account for their correlation. The correlation is kept also when aggregating the same sectors across countries and when aggregating subcategories, with some exceptions and caveats detailed in the main text.

To summarise:
- global CO₂ emitted from the energy sector alone (IPCC sector 1) accounts for 96% of global GHG emissions, and is accurate within 7% (generally, high confidence levels for top emitters);
- when adding CH₄ and N₂O, the accuracy of the energy sector decreases to an uncertainty of -12.8 to +15.9%;
- the uncertainty of N₂O for the power industry sector (factor of 10 suggested by IPCC) indicates a very poor accuracy. This high value reflects the paucity of accurate estimates, although some studies suggest lower uncertainty values (Lee et al., 2013; Olivier et al., 1999). For N₂O in sector 1, a gas set u₀ = ±50 to ±150% (industrialized and developing countries, respectively), to yield a global uncertainty of ≈12%. CH₄ emitted by the oil and gas extraction facilities is highly uncertain although the guidelines provide detailed uncertainties for all stages (extraction, storage, distribution, transmission) and differentiated by the level of development of the country. Due to the discrepancies with scientific literature and the number of parameters and components of this sector we have tested a more conservative estimate of u₀ = ±5 and ±15% (industrialised and developing countries, respectively) and u₀ = ±100% to all countries (u₀ of 50% for country specific EF) when considering aggregation of sectors/countries which yield a global CH₄ uncertainty of -55%, +93%.
- agriculture emissions are dominated by CH₄ and N₂O, with the uncertainty of the latter (over 300% on a global average) outweighing that of CH₄ due to large uncertainty in FFs. At the global scale, CH₄ uncertainty is driven by rice cultivation and enteric fermentation;
- waste is also a sector dominated by CH₄ emissions, followed by N₂O. The uncertainty of the latter are very high (often exceeding 400%), while for CH₄ emissions, the share from landfills and domestic wastewater is approximately equivalent (~44% and ~41%, respectively), whilst landfills have a relatively larger weight in the global uncertainty share (~55% and ~41%, respectively).

The strongest assumption, made also in previous studies, is the full correlation of subcategories and countries which introduces a further source of uncertainty – methodological uncertainty – that is very impactful. Uncertainty around methodological choices arises when there are different views about what constitutes the ‘correct’ approach for optimum decision making. This form of uncertainty might be dealt with by agreeing on a ‘reference case’ or on a list of methodological choices to allow comparability between different inventories.

The choice of methods can have a profound impact on the overall uncertainty assessment and needs to be taken into consideration when comparing inventories. For EU27-UK, for example, the choice to assume or not correlation among countries can result in a 4-fold variability of the uncertainty (4% vs 15%).

**ACKNOWLEDGEMENTS**

The authors are thankful to Dr P. Bergamaschi (JRC) for his insightful suggestions and critical review.

**FINANCIAL SUPPORT**

This research has been supported by the European Commission, Horizon 2020 Framework Programme (VERIFY, grant no. 776810). Margarita Choulga was funded by the CO₂ Human Emissions (CHE) project which received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 776186.

**DATA AVAILABILITY**
The database underlying the analysis is EDGARv5.0, it’s open access and available at: https://edgar.jrc.ec.europa.eu/overview.php?v=50_GHG, last access: 15 January 2021.

AUTHOR CONTRIBUTION

E. Solazzo: design of the study, analysis, writing; M. Crippa, D. Guizzardi, M. Muntean: emission database; M. Choulga: support in the uncertainty analysis of CO2; G. Janssens-Maenhout: emission database and design of the study.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

REFERENCES


Choulga, M, and et al., 2020. Global anthropogenic CO2 emissions and uncertainties as prior for Earth system modelling and data assimilation. ESSD Discussion, https://doi.org/10.5194/essd-2020-68


CCRI: China Cement Research Institute (www.ccement.com), 2019 (last access: June 2020)


GGFR/NOAA: 1994-2018 data for gas consumption for flaring, 2019


https://doi.org/10.5194/essd-12-1-2020


doi:10.1371/journal.pone.0050950


TABLE 1. AD UNCERTAINTY (UPPER AND LOWER LIMITS DEFINE THE 95% CI OF A NORMAL DISTRIBUTION). WHEN TWO VALUES ARE LISTED, THE SMALLER RANGE APPLIES TO INDUSTRIALISED COUNTRIES, THE LARGER RANGE TO DEVELOPING COUNTRIES

TABLE 2. CO₂ UNCERTAINTY OF EF BY FUEL-TYPE (FROM TABLE 3.2.1 OF IPCC-06)

TABLE 3. UNCERTAINTY OF EF FOR CH₄ AND N₂O DEFINED BY IPCC-06 SECTORS AND CORRESPONDING EDGAR SECTORS

TABLE 4. SHARE OF GHG EMISSIONS (DERIVED FROM CO₂, CH₄ AND N₂O EXPRESSED IN CO₂EQ) OF DEVELOPING AND INDUSTRIALISED COUNTRIES FOR SECTOR 1.A BASED ON EDGAR EMISSIONS FOR THE YEAR 2015

TABLE 5. SHARE OF CH₄ EMISSION IN SECTOR 1.E2B (FUGITIVE EMISSIONS FROM NATURAL GAS) FOR THE FIVE TOP EMITTING COUNTRIES

ha spostato (inserimento) [1]
ha spostato in basso [2]: COUNTRY CODES, NAMES AND DEVELOPMENT STATUS
TABLE 6. GLOBAL GHG EMISSIONS (YEAR 2015) BY IPCC SECTORS AND UNCERTAINTY RANGES DEFINING THE 95% CI OF A LOGNORMAL DISTRIBUTION

TABLE 5. COUNTRY CODES, NAMES AND DEVELOPMENT STATUS

FIGURES

FIGURE 1. GHG EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 1.A (ENERGY FROM FUEL COMBUSTION). A) CO2 FROM ENERGY INDUSTRIES; B) CO2 FROM MANUFACTURING INDUSTRIES; C) CO2 FROM TRANSPORT; D) CH4 FROM FUEL COMBUSTION; E) N2O FROM FUEL COMBUSTION; F) WORLD TOTAL. TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY’S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%), MEDIUM-HIGH (10,20%), MEDIUM (20,40%), MEDIUM-LOW (40,60%), LOW (60,100%), VERY LOW > 100% (COUNTRY CODES ARE EXPLICATED IN TABLE S1).

FIGURE 2. CO2 UNCERTAINTY AND EMISSIONS SHARES FOR EDGAR EMISSION SECTORS UNDER IPCC CATEGORY 1.A FOR BRAZIL, CHINA, GERMANY, INDIA, JAPAN, RUSSIA, SAUDI ARABIA, UNITED STATES OF AMERICA.

FIGURE 3. A) N2O EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 1.A (ENERGY FROM FUEL COMBUSTION) WHEN UNCERTAINTIES ARE SET IN THE RANGE ±50% (INDUSTRIALISED COUNTRIES) TO 150% (DEVELOPING COUNTRIES) B) WORLD TOTAL. TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY’S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%), MEDIUM-HIGH (10,20%), MEDIUM (20,40%), MEDIUM-LOW (40,60%), LOW (60,100%), VERY LOW > 100% (COUNTRY CODES ARE EXPLICATED IN TABLE S1).

FIGURE 4: GHG EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 1.B (ENERGY - FUGITIVE EMISSIONS). A) CO2 FROM FUGITIVE EMISSIONS FROM FUELS; B) CH4 FROM FUGITIVE EMISSIONS FROM SOLID FUELS; C) CH4 FROM FUGITIVE EMISSIONS FROM OIL AND NATURAL GAS; D) N2O FROM FUGITIVE EMISSIONS FROM OIL AND NATURAL GAS; E) WORLD TOTAL. TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY’S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%), MEDIUM-HIGH (10,20%), MEDIUM (20,40%), MEDIUM-LOW (40,60%), LOW (60,100%), VERY LOW > 100% (COUNTRY CODES ARE EXPLICATED IN TABLE S1).

FIGURE 6. GHG EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 2 (INDUSTRIAL PROCESSES AND PRODUCT USE): A) CO2; B) CH4; C) N2O; D) WORLD TOTAL: TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY'S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%); MEDIUM-HIGH (10,20%); MEDIUM (20,40%); MEDIUM-LOW (40,60%); LOW (60,100%). VERY LOW > 100% (COUNTRY CODES ARE EXPLAINED IN TABLE 5).

FIGURE 7. GHG EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 3 (AGRICULTURE) IN CO2 EQ (T/Year): A) CO2; B) CH4; C) N2O; D) WORLD TOTAL: TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY'S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%); MEDIUM-HIGH (10,20%); MEDIUM (20,40%); MEDIUM-LOW (40,60%); LOW (60,100%). VERY LOW > 100% (COUNTRY CODES ARE EXPLAINED IN TABLE 5).

FIGURE 8. CH4 UNCERTAINTY AND EMISSIONS SHARES FOR EDGAR'S EMISSION SECTORS UNDER IPCC CATEGORY 3 FOR BRAZIL, CHINA, INDOONESIA, INDIA, MEXICO, RUSSIA, UNITED STATES OF AMERICA, AND THE WORLD.

FIGURE 9. GHG EMISSIONS FROM TOP EMITTERS AND WORLD FOR SECTOR 4 (WASTE): A) CO2; B) CH4; C) N2O; D) WORLD TOTAL: TOTAL UNCERTAINTY; EMISSION AND UNCERTAINTY SHARES. COUNTRY'S NAMES ARE COLOR-CODED ACCORDING TO THEIR CLASSIFICATION (CYAN: INDUSTRIALISED; RED: DEVELOPING). CONFIDENCE LEVELS ARE GIVEN IN THE RANGES: HIGH (0,10%); MEDIUM-HIGH (10,20%); MEDIUM (20,40%); MEDIUM-LOW (40,60%); LOW (60,100%). VERY LOW > 100% (COUNTRY CODES ARE EXPLAINED IN TABLE 5).

FIGURE 10. CH4 UNCERTAINTY AND EMISSIONS SHARES FOR EDGAR'S EMISSION SECTORS UNDER IPCC CATEGORY 4 FOR BRAZIL, CHINA, INDOONESIA, INDIA, MEXICO, RUSSIA, UNITED STATES OF AMERICA, AND THE WORLD.

FIGURE 11. GLOBAL SHARE OF EMISSIONS AND UNCERTAINTY BY A) GAS AND B) CATEGORY

FIGURE 12. EU27+UK SHARE OF EMISSIONS AND UNCERTAINTY BY A) GAS AND B) CATEGORY

FIGURE 13. VARIABILITY OF WORLD EMISSIONS UNCERTAINTY INTRODUCED BY METHODOLOGICAL CHOICE: TIER 1+CORREL. (IN RED): IS THE BASE CASE AND ASSUMES CORRELATION AMONG ALL SECTORS AND AMONG SECTORS ACROSS COUNTRIES AND DEFAULT TIER 1 IPCC 95% UNCERTAINTY. TIER1+CORREL. (IN GREEN) DIFFERS FROM THE BASE CASE ONLY FOR THE N2O UNCERTAINTY IN SECTOR 1A (±100% TO ±150%). (GREEN) TIER1+CORREL. (IN BLUE) DIFFERS FROM THE BASE CASE AS IT ASSUMES NO CORRELATION AND NO UNCERTAINTY IN SECTOR 1A, (IN THE RANGE) ±100% TO ±150%.

FIGURE 14. VARIABILITY OF EU27+UK EMISSIONS UNCERTAINTY INTRODUCED BY METHODOLOGICAL CHOICE: TIER 1+CORREL. (IN RED): IS THE BASE CASE AND ASSUMES CORRELATION AMONG ALL SECTORS AND AMONG SECTORS ACROSS COUNTRIES AND DEFAULT TIER 1 IPCC 95% UNCERTAINTY. TIER1+CORREL. (IN GREEN) DIFFERS FROM THE BASE CASE ONLY FOR THE N2O UNCERTAINTY IN SECTOR 1A (±100% TO ±150%). (GREEN) TIER1+CORREL. (IN BLUE) DIFFERS FROM THE BASE CASE AS IT ASSUMES NO CORRELATION AND NO UNCERTAINTY IN SECTOR 1A, (IN THE RANGE) ±100% TO ±150%.