Evaluating Asymmetric Approaches to the Estimation of Standard Uncertainties for Emission Factors in the Fuel Sector of Costa Rica

Gabriel Molina-Castro* and Bryan Calderón-Jiménez

Chemical Metrology Division, National Metrology Laboratory of Costa Rica, San Jose, Costa Rica

A new environmental challenge for Costa Rica involves the precise and reliable quantification of data from its fossil-fueled transportation sector. In the context of greenhouse gas inventories (measurement), uncertainty assessment, as the best quality parameter of any estimation or measurement, takes on a new relevance by becoming a mandatory requirement on ISO 14064-1:2018. However, a significant limitation has been found by users when quantifying standard (measurement) uncertainties associated with emission factors with asymmetric probability distributions. The present article sought to take advantage of fitting asymmetric distributions to estimate and compare possible standard uncertainties for the official emission factors of Costa Rica, specifically for the fuel sector. Five asymmetric distributions and a “symmetrization” method (symmetric approximation of an asymmetric distribution) were chosen and fitted to the data based on their application and previous use. Standard uncertainties were estimated from each distribution parameters as standard deviations. To evaluate the fit, quantiles of interest were extracted from simulated populations compared with the original data values. A systematically better fit was evidenced for the asymmetric triangular and generalized extreme value distributions, both for CO₂ emission factors with less asymmetries and CH₄ and N₂O emission factors with greater asymmetries. This was not the case for the other distributions, where the log-normal distribution applying the correction factor suggested in the literature showed the worst fit. The use of the former distributions is recommended to estimate the standard uncertainties associated with the emission factors from the official Costa Rican database and other emission factors with similar asymmetries.

Keywords: asymmetric distribution, emission factor, greenhouse gases inventories, uncertainty, fuel emissions

INTRODUCTION

Costa Rica is recognized worldwide as a pioneer in environmental protection and the fight against climate change (UN Environment Programme, 2019). Its efforts to achieve high recovery of its forest cover (Allen and Padgett-Vásquez, 2017), ecotourism as a key component of its economy (Courvisanos and Ameeta, 2006) and an electricity grid that is almost 100 % renewable (ICE Group, 2020) stand out. The new environmental challenge for Costa Rica involves its fossil-fueled...
transformation sector that generates large amounts of emissions and a variety of other problems (Fendt, 2017). In recent years, the government of Costa Rica has launched national policies and plans that seek to decarbonize its economy (Government of Costa Rica, 2019), including its current transportation system and the environmental impact it generates. To achieve the success of these policies, it is essential to have precise and reliable data that allow the correct quantification of measures taken and guarantee the success of environmental efforts in favor of its conservation.

Measurement uncertainty, referred in this study simply as “uncertainty” and defined as the doubt about the true value of a quantity that remains after making that measurement or estimation, is the best quality parameter of any measurement or estimation and reflects the impossibility of knowing exactly its value (Joint Committee for Guides in Metrology [JCGM], 2008). Among the accepted methodologies for estimating uncertainty, the application of the law of propagation of uncertainty or Gauss’s formula included in the Guide to the Expression of Uncertainty in Measurement (GUM) stands out. This methodology is based on modeling an output quantity \( y \) as a known function of several input quantities and handles the uncertainties associated with the input quantities by modeling them as random variables. This approach uses the standard deviations of these random variables and their correlations to produce an approximate evaluation of the standard (measurement) uncertainty \( u_y \), a measurement uncertainty expressed as a standard deviation associated to the output quantity \( y \) (Joint Committee for Guides in Metrology [JCGM], 2012). It should be highlighted that the conventional technique described in the GUM does not require necessarily that probability distributions of the input quantities to be symmetrical (Possolo et al., 2019).

In the context of greenhouse gas (GHG) inventories, uncertainty evaluation was initially considered as a non-mandatory requirement by ISO 14064-1 (2006), the main international standard for GHG quantification and reporting by organizations. Thus, the implementation of uncertainty estimates in GHG inventories has not yet been covered by many organizations that have already quantified emissions and mitigated their impact on the environment, depriving them of the measure of quality that strengthens the confidence of their results. Nevertheless, uncertainty in GHG inventories has been pointed as a key component to consider (Rypdal and Winiwater, 2001; EPA, 2002; Jonas et al., 2010) and several studies have been developed regarding this topic, including EPA (1996), Ritter et al. (2010), Milne et al. (2015), and Solazzo et al. (2020). With the publication of the new version of ISO 14064-1 (2018), uncertainty assessment takes on a new dimension by becoming a mandatory requirement. This conjuncture added to Costa Rican national scaled efforts to include uncertainty assessment through the National Carbon Neutrality Program (PPCN, by its Spanish acronym), has aroused national interest in the application of uncertainty estimation methodologies in GHG quantification (DCC and PMR, 2020; DCC and LCM, 2020).

A common scenario found in GHG inventories is the indirect quantification of emissions, where emissions are not measured directly as an amount of gas released into the atmosphere. Instead, emissions \( E \) are estimated from other data values associated with the activity that cause the emission \( d \) and emission factors \( f \) that relate these data to the amount of gas emitted, as shown in Equation (1). Emissions from energy consumption are a typical example of indirect quantification, where the emissions from e.g., a furnace can be estimated from the amount of fuel consumed and an emission factor that relates the liters of fuel with the amount of GHG released by the combustion process.

\[
E = d \cdot f
\]  

(1)

Following the application of Gauss’s formula (Joint Committee for Guides in Metrology [JCGM], 2008; Possolo and Iyer, 2017) to Equation (1), the approximate standard uncertainty of \( E (u_E) \) can be easily estimated from the values and standard uncertainties of \( d (u_d) \) and \( f (u_f) \), according to Equation (2).

\[
u_E = \sqrt{\left(\frac{u_d}{d}\right)^2 + \left(\frac{u_f}{f}\right)^2}
\]  

(2)

However, a significant limitation has been found by users when quantifying \( u_f \) since the literature usually includes references to asymmetric ranges of variation when describing the behavior of emission factors (Intergovernmental Panel on Climate Change [IPCC], 2000, 2006), including the official database of Costa Rican emission factors (IMN, 2020). This is attributable to the observed behavior when studying these factors, which usually present important dispersions depending on the study conditions. Small values are frequently found but relatively high values are plausible too, making asymmetry the commonly encountered behavior (Bharvirkar, 1999; Frey, 2007). Furthermore, the problem of propagating uncertainties expressed asymmetrically is not addressed in the GUM and has been treated on a limited basis by Barlow (2004) and Audi et al. (2017) using “symmetrized” approximations and by Possolo et al. (2019) adjusting asymmetric distributions, among others. In the specific case of GHG emission factors, some guidelines published by the Intergovernmental Panel on Climate Change (IPCC) briefly address the issue (Intergovernmental Panel on Climate Change [IPCC], 2000, 2006), although they do not delve into the methodologies for estimating standard uncertainties for these asymmetric ranges.

The present article sought to take advantage of fitting asymmetric distributions to estimate and compare possible standard uncertainties for the official emission factors of Costa Rica, specifically for the fuel sector. The choice of this sector was due to its strategic importance in Costa Rican national environmental policies, and the fact that it represents the largest group of factors and contains the greatest asymmetries within the entire official database. It should be noted that this study addresses the uncertainties of the emission factors as standard uncertainties (standard deviations) according to the GUM guidelines, not as probability ranges or confidence intervals as is the typical approach on the topic (Intergovernmental Panel on Climate Change [IPCC], 2006; Choulga et al., 2020; Solazzo et al., 2020). It is expected that this study will serve as a guide for the interpretation and manipulation of uncertainties within the process of implementing uncertainty estimation in GHG...
inventories, helping to obtain a more reliable quantification of emission data from fuels in Costa Rica and other countries around the world.

MATERIALS AND METHODS

Five asymmetric distributions were chosen based on their application and previous use in related technical documents including Intergovernmental Panel on Climate Change [IPCC] (2000, 2006), Possolo et al. (2019), and DCC and LCM (2020): the asymmetric triangular distribution, the log-normal distribution, the Fechner distribution, the skew-normal distribution and generalized extreme value distribution. An example of these distributions is shown in Figure 1.

These distributions were fitted to the official database of Costa Rican emission factors for the fuel sector (IMN, 2020). This database includes the accepted value for the emission factor \( f \) and two additional values \( U_R \) (also called “right uncertainty” or “upper uncertainty”, corresponding to the upper bound of the 95% probability range estimated for the emission factor) and \( U_L \) (also called “left uncertainty” or “lower uncertainty,” corresponding to the lower bound of the 95% probability range estimated for the emission factor) that delimit an asymmetric interval of possible values. Subsequently, this information was used to estimate the parameters of each distribution in order to calculate its theoretical variance \( \sigma^2 \) and its corresponding standard uncertainty \( u_X = \sqrt{\sigma^2} \).

The details of each distribution and the chosen symmetrization method are shown below. Except in cases where it is specified, the distribution mean \( \mu \) was set to \( f \).

Following the methodology proposed by Possolo et al. (2019), it was considered that each input \( f \) has an underlying asymmetric probability distribution qualified by \( U_R \) and \( U_L \), both positives. However, according to IPCC good practices and guidelines (Intergovernmental Panel on Climate Change [IPCC], 2000, 2006), the interval \([ f - U_L, f + U_R] \) covers the true value of \( f \) with an approximate probability of 95%. Also, it was assumed that “probabilities \( 0 < p_L < p_M < p_R < 1 \) are specified such that \( p_L = Pr\{X < f - U_L\}, p_M = Pr\{X < f\}, \) and \( p_R = Pr\{X < f + U_R\} \), where \( X \) denotes the random variable modeling the uncertainty associated with \( f \)” (Possolo et al., 2019). In all cases, \( p_L \) and \( p_R \) were set to 0.025 and 0.975, respectively. For the cases of Fechner, skew-normal and generalized extreme values, \( p_M \) was set to 0.50 and the applied methods sought to best reproduce \( f - U_L, f \), and \( f + U_R \) as the specified percentiles of the corresponding distribution, using a non-linear, unweighted least-squares method.

Finally, to evaluate the fit of each distribution, populations of size \( 10^6 \) were simulated with the estimated uncertainties or the distribution parameters that originated them. Since the present study focuses on uncertainty estimation, quantiles 2.5% \( (q_{0.025}) \) and 97.5% \( (q_{0.975}) \) were extracted from each generated population and compared with the original values of \( U_L \) and \( U_R \) included in the official database. The largest absolute relative error \( (RE) \) was reported as fitting statistic, according to Equation (3).

\[
RE = \max \left( \left| \frac{q_{0.025} - U_L}{U_L} \right|, \left| \frac{q_{0.975} - U_R}{U_R} \right| \right)
\]

For all the calculations, data processing, statistical evaluation and simulations, the free environment for statistical computing R version 3.6.1 (R Core Team, 2020) was used. The programming code used in this study can be consulted in Supplementary Material.

Asymmetric Triangular Distribution \((u_{Tri})\)

As its name indicates, this distribution consists of a triangle whose maximum height is not in its center, and can be asymmetric to the left or to the right. The asymmetric triangular distribution consists of three parameters: the mode \( \mu \), the lower extreme value \( a \), and the upper extreme value \( b \). The variance of any triangular distribution is theoretically estimated by following Equation (4).

\[
\sigma^2 = \frac{a^2 + \mu^2 + b^2 - a\mu - b\mu + ab}{18}
\]

For the estimation of the parameters, the geometric estimation of the two-half triangles area was used (Petty and Dye, 2013). Considering \( \mu = f \), the equation system shown in (5) was obtained. This system was solved using Broyden and Newton methods (Dennis and Schnabel, 1996) included in the R-package \texttt{nleqslv} functions (Hasselman, 2018). For the population simulation, computational facilities provided by R-package \texttt{triangle} (Carnell, 2019) were used.

\[
\begin{align*}
\left\{ \begin{array}{l}
b - a - (U_L - a)^2 / (0.025 (f - a)) \\
b - a - (b - U_R)^2 / (0.025 (b - f)) 
\end{array} \right.
\end{align*}
\]

Log-Normal Distribution \((u_{LN})\)

As its name indicates, the logarithmic normal distribution corresponds to the probability distribution of a variable whose natural logarithm results in a normal distribution. Its most common uses include modeling the multiplicative products of independent factors and strictly positive continuous variables with wide ranges (Intergovernmental Panel on Climate Change [IPCC], 2006). This distribution consists of two parameters: the mean \( \mu \) and the variance \( \sigma^2 \), corresponding to the natural logarithm of the mean and variance of the normally distributed variable. According to the IPCC guidelines, it is possible to estimate the distribution variance \( \sigma^2 \) from uncertainties \( U_L, U_R \) and mean \( \mu \) following Equation (6), where \( \sigma_g \) corresponds to the geometric standard deviation of the distribution, estimated according to Equation (7).

\[
\sigma^2 = \mu^2 \cdot e^{(\ln \sigma_g)^2} - 1
\]

\[
-1.96 \cdot \ln \sigma_g = \frac{(\ln U_L + \ln U_R)}{2}
\]

Finally, a correction factor \( F_C \) is recommended to slightly increase \( u \) as \( u_{LNC} = u_{LN} \cdot F_C \) for cases with high standard
uncertainties ($u_{rel} > 50\%$) (Intergovernmental Panel on Climate Change [IPCC], 2006; DCC and LCM, 2020). The estimation of the factor $F_C$ is shown in Equation (8), adapted from Frey (2003) and Intergovernmental Panel on Climate Change [IPCC] (2006) considering the use of a standard uncertainty ($u$) instead of a half interval uncertainty ($U$), where $u_{rel} = 100 \cdot u_{LN}/f$. Both $u_{LN}$ and $u_{LNC}$ were estimated and evaluated in the present study. For the simulation process, R-package stats functions were used (R Core Team, 2020).

The Fechner distribution consists of three parameters: the mode $\mu$, the left variance $\sigma_L^2$ and the right variance $\sigma_R^2$. The Fechner distribution variance is theoretically estimated following Equation (9). A Fechner distribution may be a suitable model if $0.410 < U_R/ U_L < 2.44$ for a coverage probability of 95 % (Possolo et al., 2019).

$$\sigma^2 = \left( 1 - \frac{2}{\pi} \right) (\sigma_R - \sigma_L)^2 + \sigma_R \sigma_L$$

(9)

For the estimation of the parameters and the simulation process, the code already generated by Possolo et al. (2019) was used, which uses the computational facilities of R-package fanplot (Abel, 2015).

$\textbf{Skew-Normal Distribution ($u_{SN}$)}$

This distribution is a generalization of the Gaussian (normal) distribution that considers an additional bias parameter ($\alpha$). This bias allows this distribution for asymmetry to the left or right, depending on the value of $\alpha$ (for the traditional Gaussian distribution, $\alpha = 0$). The skew-normal distribution consists of three parameters: a location parameter $\mu$, a scale parameter $\omega > 0$, and the bias parameter $\alpha$ (Azzalini and Capitanio, 2014).

The skew-normal distribution variance is theoretically estimated by following Equation (10). A skew-normal distribution may be a

$$\sigma^2 = \frac{1}{\sqrt{2\pi}} \left( 1 + \frac{1}{2} \alpha^2 \right) \sigma\omega$$

(10)
suitable model if $0.410 < U_R/U_L < 2.44$ for a coverage probability of 95% (Possolo et al., 2019).

\[
\sigma^2 = \frac{\omega^2}{\pi} \left[1 - \frac{2\omega^2}{\pi (1 + \omega^2)}\right] \quad (10)
\]

For the estimation of the parameters and the simulation process, the code already generated by Possolo et al. (2019) was used, which uses the computational facilities of R-package evd (Azzalini, 2020).

**Generalized Extreme Value Distribution**

The extreme value distributions are generally considered to comprise three distribution families, including the Gumbel, Fréchet, Gompertz, and reverse Weibull distributions (Possolo et al., 2019). These distributions may all be represented as members of a single family of generalized distributions with a common cumulative distribution function, known as the generalized extreme value distribution (GEVD). The GEVD consists of three parameters: a location parameter $\mu$, a scale parameter $\omega > 0$, and a shape parameter $\xi$ (Johnson et al., 1995). The GEVD variance is theoretically estimated by following Equation (11),

\[
\sigma^2 = \left[g_2 - g_1^2\right] \frac{\omega^2}{\xi^2} \quad (11)
\]

where $g_k = \Gamma \left(1 - k\xi\right)$ and $\Gamma$ is the gamma function. For the estimation of the parameters and the simulation process, the code already generated by Possolo et al. (2019) was used, which uses the computational facilities of R-package evd (Stephenson, 2002).

**Symmetrization Method**

This method consists of a simple approximation of an asymmetric distribution as a "symmetric" normal distribution. Audi et al. (2017) mention it as method 1 and only requires estimating the difference between $U_R$ and $U_L$. The resulting value is considered equivalent to the coverage interval (with the same coverage probability as interval $[f - U_L,f + U_R]$) for a normal distribution centered at $\mu = (U_R + U_L)/2$. Considering the properties of the normal distribution, it is possible to obtain the relationship shown in Equation (12), where $\sigma$ is the standard deviation for the "symmetrized" distribution. For the simulation process, R-package stats functions were used (R Core Team, 2020).

\[
\sigma = \frac{U_L - U_R}{4} \quad (12)
\]

**RESULTS AND DISCUSSION**

Figure 2 shows an example of the obtained results for each of the asymmetric distributions and the symmetrization method applied to a single source: lubricant (for commercial, institutional, residential, agricultural and land transportation use). This specific source was selected because its results are highly representative of the general behavior observed in the other sources and factors. Rows on the figure represent the three emission factors (CO$_2$, CH$_4$, and N$_2$O) included in the official Costa Rican database for most sources, including lubricant. Upper side of the figure illustrates the simulated distributions behavior and lower side of the figure compares the estimated intervals defined by quantiles 2.5 % ($q_{2.5}$) and 97.5 % ($q_{97.5}$) of the simulated populations with the $[f - U_L,f + U_R]$ reference interval as a visual aid to interpret the RE values.

Tables 1–3 show the complete results for each of the asymmetric distributions and the symmetrization method applied to all the official Costa Rican emission factors for the fuel sector. These results correspond to the estimated standard uncertainties ($u$) and the maximum relative errors obtained for the quantiles of interest (RE) after the respective simulations and intervals estimation. All estimated standard uncertainties are reported as absolute uncertainties. However, due to their widespread use in the GHG sector, the corresponding relative standard uncertainties are shown in Supplementary Tables 1–3.

Table 1 and Supplementary Table 1 show the results corresponding to the CO$_2$ emission factors (left column of Figure 2 as example). These factors have the particularity of reporting the least asymmetric intervals of all the analyzed database. All distributions have standard uncertainty values similar to each other, except for the corrected log-normal distribution. This behavior was expected since correction factor $F_C$ is only recommended for cases with high standard uncertainties. The asymmetric triangular distribution systematically showed the largest standard uncertainties, while the smallest standard uncertainties were evidenced with the corrected log-normal distribution or the symmetrization method. The greatest differences between the standard uncertainties were evidenced in the emission factor for airplane gasoline, corresponding to the most asymmetric interval in the tables. It should be noted that the uncertainties estimated for this factor with Fechner ($u_{\text{Fech}}$) and skew-normal ($u_{\text{SN}}$) distributions may not be suitable for the interval asymmetry, due to the non-compliance of the recommendation $0.410 < U_R/U_L < 2.44$ in both cases according to Possolo et al. (2019).

When evaluating the distribution fitting in Table 1 and Supplementary Table 1 through RE values for the simulated populations, it is noted that the best fitting is obtained for the asymmetric triangular and GEV distributions with $RE < 0.03$ in all cases. Subsequently, the Fechner and skew-normal distributions showed $RE < 0.05$ in all cases except for the airplane gasoline factor mentioned above ($RE_{\text{Fech}} = 0.58$, $RE_{\text{SN}} = 0.59$). The symmetrization method and the log-normal distribution showed similar fittings, with a slight improvement for the first ($RE_{\text{Sym}} \leq 0.35$). Finally, the corrected log-normal distribution did not show to be useful for such small asymmetries and small relative uncertainty values as expected ($RE_{\text{LNC}} \leq 7.11$).

Table 2 and Supplementary Table 2 show the results corresponding to the CH$_4$ emission factors (middle column of Figure 2 as example) while Table 3 and Supplementary Table 3 do the same for N$_2$O emission factors (right column of Figure 2 as example). The factors for both gases showed percentage variation intervals with asymmetries greater than...
### TABLE 1 | Absolute standard uncertainties $u$ and largest relative error $RE$ estimated for CO$_2$ emission factors for the fuel sector using various asymmetric approaches.

The best estimate $f$ and expanded asymmetric uncertainties $U_f$ and $U_H$ were taken from the official database of Costa Rican emission factors (IMN, 2020). Values of $RE$ are shown inside the brackets.

| Fuel | $f$ (kg/L) | $U_f$ (kg/L) | $U_H$ (kg/L) | $u$ (kg/L) [RE (%)] |
|------|------------|--------------|--------------|---------------------|
|      | Tri        | LN           | LNC          | Fech               |
|      |            |              |              | SN                 |
|      |            |              |              | GEV                |
|      |            |              |              | Sym                |
| Gasoline | 2.231 | 2.129 | 2.362 | 0.062 [0.00] | 0.059 [0.55] | 0.053 [1.05] | 0.060 [0.01] | 0.060 [0.01] | 0.058 [0.11] |
| Diesel | 2.613 | 2.531 | 2.696 | 0.043 [0.00] | 0.042 [0.01] | 0.031 [0.81] | 0.042 [0.00] | 0.042 [0.01] | 0.043 [0.01] | 0.041 [0.07] |
| Bunker | 3.101 | 2.990 | 3.214 | 0.059 [0.00] | 0.057 [0.01] | 0.045 [0.76] | 0.057 [0.00] | 0.057 [0.01] | 0.058 [0.01] | 0.056 [0.07] |
| Kerosene | 2.541 | 2.444 | 2.645 | 0.053 [0.00] | 0.051 [0.08] | 0.042 [0.77] | 0.051 [0.01] | 0.051 [0.01] | 0.052 [0.01] | 0.050 [0.08] |
| LPG | 1.611 | 1.476 | 1.759 | 0.074 [0.01] | 0.072 [0.09] | 0.072 [0.12] | 0.072 [0.01] | 0.072 [0.03] | 0.073 [0.02] | 0.071 [0.19] |
| Airplane gasoline | 2.227 | 2.050 | 2.750 | 0.188 [0.01] | 0.167 [6.48] | 0.175 [7.11] | 0.193 [0.58] | 0.193 [0.59] | 0.189 [0.01] | 0.175 [0.35] |
| Jet fuel | 2.505 | 2.388 | 2.638 | 0.066 [0.00] | 0.064 [0.23] | 0.057 [0.77] | 0.064 [0.01] | 0.064 [0.01] | 0.064 [0.01] | 0.063 [0.11] |
| Lubricant | 2.549 | 2.250 | 2.874 | 0.164 [0.01] | 0.159 [0.06] | 0.164 [0.43] | 0.159 [0.04] | 0.161 [0.02] | 0.156 [0.28] |

*Distribution may not be suitable according to Possolo et al. (2019).

### TABLE 2 | Absolute standard uncertainties $u$ and largest relative error $RE$ estimated for CH$_4$ emission factors for the fuel sector using various asymmetric approaches.

The best estimate $f$ and expanded asymmetric uncertainties $U_f$ and $U_H$ were taken from the official database of Costa Rican emission factors (IMN, 2020). Values of $RE$ are shown inside the brackets.

| Source/Fuel | $f$ (g/L) | $U_f$ (g/L) | $U_H$ (g/L) | $u$ (g/L) [RE (%)] |
|-------------|-----------|-------------|-------------|---------------------|
|             | Tri       | LN*         | LNC         | Fech*               |
|             |            |              |              | SN*                |
|             |            |              |              | GEV                |
|             |            |              |              | Sym*               |
| Electricity generation, manufacturing and construction/Diesel | 0.122 | 0.0354 | 0.3550 | 0.0858 [0.0785] | 0.0882 [0.0876] | 0.0876 [0.0848] | 0.0848 [0.0799] |
| Electricity generation, manufacturing and construction/Bunker | 0.138 | 0.0304 | 0.4016 | 0.0993 [0.1017] | 0.1204 [0.1013] | 0.1013 [0.0972] | 0.0972 [0.0928] |
| Manufacturing and construction/Gasoline | 0.111 | 0.0322 | 0.3219 | 0.0777 [0.0713] | 0.0801 [0.0794] | 0.0794 [0.0768] | 0.0768 [0.0724] |
| Manufacturing and construction/LPG | 0.027 | 0.0076 | 0.0753 | 0.0181 [0.0173] | 0.0194 [0.0185] | 0.0185 [0.0178] | 0.0178 [0.0169] |
| Manufacturing and construction/Lubricant | 0.104 | 0.0281 | 0.2881 | 0.0695 [0.0676] | 0.0763 [0.0709] | 0.0709 [0.0680] | 0.0680 [0.0650] |
| Commercial, institutional, residential and agricultural/Gasoline | 0.346 | 0.0969 | 0.9653 | 0.2324 [0.2217] | 0.2490 [0.2371] | 0.2371 [0.2279] | 0.2279 [0.2171] |
| Commercial, institutional, residential and agricultural/Diesel | 0.382 | 0.1108 | 1.0581 | 0.2535 [0.2395] | 0.2673 [0.2587] | 0.2587 [0.2487] | 0.2487 [0.2388] |
| Commercial, institutional, residential and agricultural/Bunker | 0.433 | 0.1212 | 1.1994 | 0.2884 [0.2764] | 0.3102 [0.2944] | 0.2944 [0.2826] | 0.2826 [0.2695] |
| Commercial, institutional, residential and agricultural/LPG | 0.139 | 0.0389 | 0.3878 | 0.0934 [0.0981] | 0.1000 [0.0953] | 0.0953 [0.0916] | 0.0916 [0.0872] |
| Commercial, institutional, residential, agricultural and land transportation/Lubricant | 0.348 | 0.0870 | 0.9013 | 0.2168 [0.2275] | 0.2571 [0.2192] | 0.2192 [0.2106] | 0.2106 [0.2036] |
| Land transportation/Gasoline/no catalyst | 1.176 | 0.3058 | 3.5398 | 0.8680 [0.8126] | 0.9366 [0.8872] | 0.8872 [0.8598] | 0.8598 [0.8085] |
| Land transportation/Gasoline/catalyst | 0.907 | 0.2358 | 2.7573 | 0.6771 [0.6298] | 0.7273 [0.6923] | 0.6923 [0.6721] | 0.6721 [0.6304] |
| Land transportation/Diesel/no catalyst | 0.149 | 0.0566 | 0.3367 | 0.0745 [0.0714] | 0.0758 [0.0750] | 0.0750 [0.0741] | 0.0741 [0.0722] |

*Distribution may not be suitable according to Possolo et al. (2019).

* Estimated uncertainties may need to be multiplied by $F_c$ according to Intergovernmental Panel on Climate Change (IPCC), 2006.

*Estimated uncertainties may be suitable without applying any limitation or correction.

*Not recommended due to significant number of negative values obtained in the simulations.

the CO$_2$ factors and similar to each other. The corrected log-normal distribution systematically showed the largest standard uncertainties (except for one case in N$_2$O), while the smallest standard uncertainties were evidenced for the log-normal distribution or the symmetrization method. However, the results obtained for the symmetrization method should be taken with caution. Given the characteristics of the normal distribution assumed by this method and the proximity of the intervals to 0, the simulated populations presented a significant percentage of negative values (between 5 % and 10 %), which contradict...
the physical nature of an emission factor. The asymmetric triangular, Fechner, skew-normal, and GEV distributions shared the possibility to theoretically present negative results, but the proportions of values below 0 obtained in the simulated populations were considered insignificant (<0.2 %). For this reason, only the use of this symmetrization method is not recommended for the evaluated CH₄ and N₂O emission factors.

The greatest differences between standard uncertainties for the CH₄ emission factors were evidenced for lubricants due to the little similarity between the corrected log-normal distribution and the other distributions. It should be noted that these cases do not correspond to the factor with the greatest asymmetry in Table 2 and Supplementary Table 2, as occurred with the CO₂ emission factors. For the N₂O emission factors, the greatest differences were evidenced for the gasoline without catalyst in land transportation factor. This factor does correspond to the most asymmetric interval in Table 3 and Supplementary Table 3, with reduced values for both log-normal methods while the highest value was evidenced in the GEV distribution (the only emission factor with this behavior).

It should be noted again that most of $u_{\text{Fech}}$ and $u_{\text{SN}}$ may not be suitable for the emission factor asymmetries of both gases due to non-compliance with recommendation $0.410 < U_R / U_L < 2.44$ (Possolo et al., 2019). Compliance with this criterion was only evidenced for five CH₄ emission factors and five N₂O emission factors. Uncertainties estimated with the log-normal distribution $u_{\text{LN}}$ should also be highlighted. According to the IPCC guidelines (Intergovernmental Panel on Climate Change [IPCC], 2006), these uncertainties could be underestimated due to their high percentage values relative to the emission factor ($u_{\text{rel}} > 50 \%$). Therefore, it would be initially recommended to use $u_{\text{LNC}}$ instead.

When evaluating the distribution fitting in the factors of both gases through $RE$ values for the simulated populations, it is again noted that the best fitting is obtained for the asymmetric triangular distribution ($RE_{\text{Tri}} < 0.2$) and GEV distribution ($RE_{\text{GEV}} < 0.5$), respectively. The case of N₂O emission factor for gasoline without catalyst in land transportation should be highlighted, where these methods seem to fit adequately ($RE_{\text{Tri}} = 0.05, RE_{\text{GEV}} = 0.11$), but their relative standard uncertainties differ by more than 9 % ($u_{\text{Tri}} = 68.71 \%, u_{\text{GEV}} = 78.10 \%)$. This situation does not occur in any other emission factor for CO₂, CH₄, or N₂O, where RE values and standard uncertainties between both distributions seem

### Table 3: Absolute standard uncertainties $u$ and largest relative error $RE$ estimated for N₂O emission factors for the fuel sector using various asymmetric approaches.

The best estimate $f$ and expanded asymmetric uncertainties $U_L$ and $U_R$ were taken from the official database of Costa Rican emission factors (IMN, 2020). Values of RE are shown inside the brackets.

| Source/Fuel | $f$ (g/L) | $U_L$ (g/L) | $U_R$ (g/L) | $u$ (g/L) | [Tri (g/L)] | [LN (g/L)] | [Fech (g/L)] | [SN (g/L)] | [GEV (g/L)] | [Sym (g/L)] |
|-------------|-----------|-------------|-------------|----------|-------------|------------|-------------|------------|------------|------------|
| Electricity, manufacturing and construction/Diesel | 0.024 | 0.007 | 0.070 | 0.0171 | 0.0157 | 0.0176 | 0.0175 | 0.0175 | 0.0169 | 0.0159 |
| Electricity, manufacturing and construction/Bunker | 0.027 | 0.008 | 0.080 | 0.0194 | 0.0178 | 0.0200 | 0.0198 | 0.0198 | 0.0192 | 0.0181 |
| Manufacturing and construction/Gasoline | 0.022 | 0.006 | 0.064 | 0.0155 | 0.0142 | 0.0160 | 0.0158 | 0.0158 | 0.0153 | 0.0144 |
| Manufacturing and construction/LPG | 0.002 | 0.000 | 0.007 | 0.0018 | 0.0018 | 0.0020 | 0.0019 | 0.0019 | 0.0018 | 0.0017 |
| Manufacturing, construction/Lubricant | 0.021 | 0.005 | 0.058 | 0.0140 | 0.0137 | 0.0154 | 0.0143 | 0.0143 | 0.0137 | 0.0131 |
| Commercial, institutional, residential and agricultural/Gasoline | 0.022 | 0.006 | 0.064 | 0.0155 | 0.0142 | 0.0160 | 0.0158 | 0.0158 | 0.0153 | 0.0144 |
| Commercial, institutional, residential and agricultural/Diesel | 0.024 | 0.007 | 0.070 | 0.0171 | 0.0157 | 0.0176 | 0.0175 | 0.0175 | 0.0169 | 0.0159 |
| Commercial, institutional, residential and agricultural/Bunker | 0.027 | 0.008 | 0.080 | 0.0194 | 0.0178 | 0.0200 | 0.0198 | 0.0198 | 0.0192 | 0.0181 |
| Commercial, institutional, residential and agricultural/LPG | 0.002 | 0.000 | 0.007 | 0.0018 | 0.0018 | 0.0020 | 0.0019 | 0.0019 | 0.0018 | 0.0017 |
| Commercial, institutional, residential, agricultural and land transportation/Lubricant | 0.021 | 0.005 | 0.058 | 0.0140 | 0.0137 | 0.0154 | 0.0143 | 0.0143 | 0.0137 | 0.0131 |
| Land transportation/Gasoline/no catalyst | 0.116 | 0.060 | 0.352 | 0.0797 | 0.0650 | 0.0858 | 0.0819 | 0.0818 | 0.0906 | 0.0731 |
| Land transportation/Gasoline/catalyst | 0.293 | 0.082 | 0.772 | 0.1847 | 0.1761 | 0.1961 | 0.1883 | 0.1884 | 0.1808 | 0.1726 |
| Land transportation/Diesel/no catalyst | 0.154 | 0.046 | 0.423 | 0.1010 | 0.0945 | 0.1048 | 0.1031 | 0.1031 | 0.0991 | 0.0943 |
to be consistent with each other. For Fechner and skew-normal distributions, the ten emission factors in which their use may be appropriate showed very good fittings as well ($RE_{\text{Fech}} < 0.1$, $RE_{\text{SN}} \leq 0.5$), while mixed fittings were evidenced for the other factors ($0.6 < RE < 21.9$) and for the log-normal distribution ($0.6 < RE_{\text{LN}} < 28.2$). Although not recommended, the symmetrization method was not very effective when evaluating its fitting in these emission factors ($4.8 < RE_{\text{Sym}} < 14.8$). It is noteworthy that, according to the IPCC guidelines (Intergovernmental Panel on Climate Change [IPCC], 2006), the high uncertainties would justify the use of the correction factor $F_C$ to improve the estimation of the log-normal distribution. However, the corrected log-normal distribution fittings were the worst for all factors for both analyses ($8.4 < RE_{\text{LNC}} < 32.3$). In this context, a deeper analysis is recommended on the use of Equation (8) adaptation for the correction of standard uncertainties from a log-normal distribution instead of the original equation stated on Frey (2003) and Intergovernmental Panel on Climate Change [IPCC] (2006) for half interval uncertainties.

Although most of the proposed methods can be used to estimate the standard uncertainties of the emission factors included in the official Costa Rican database for the fuel sector (except for symmetrization method for CH$_4$ and N$_2$O emission factors), the asymmetric triangular and GEV distributions obtained the best fittings in the present study. For practical purposes, any of these two distributions could be chosen to estimate the standard uncertainty for Costa Rican fuel emission factors, with little difference in the results in most cases. Its use can also be recommended to address the estimation of standard uncertainties for other emission factors, even for other countries, as long as the probability of obtaining negative values is not a concern (for example, emission factors with $U_L$ values close to 0). In the latter cases, the use of a log-normal distribution is recommended or exploring the use of truncated variants of the distributions evaluated in the present study that prevent the presence of negative values.

**CONCLUSION**

From the present study, the applicability of different approaches of asymmetric distributions to address standard uncertainty estimation of emission factors in GHG inventories according to GUM guidelines was evident, specifically for the Costa Rican official database of these factors in the fuel sector. The comparability of the different methods applied was also appreciable, with significant differences in the fittings obtained for factors with probability ranges or coverage intervals reported with greater asymmetries. Overall, a systematically better fit was evidenced for the asymmetric triangular and generalized extreme value distributions, both for CO$_2$ emission factors with less asymmetries and CH$_4$ and N$_2$O emission factors with greater asymmetries. The observed fit was not as good...
for the other distributions, with the log-normal distribution applying the correction factor suggested in the literature showed the worst results.

Therefore, the use of the asymmetric triangular or generalized extreme value distributions is recommended indistinctly to estimate the standard uncertainties associated with the emission factors from the official Costa Rican database for the fuel sector. Its use is also recommended within the process of implementing uncertainty estimation in GHG inventories and improving the quality of its results, considering other emission factors for Costa Rica or other countries as well, provided that their intervals of variation are not very close to 0. Further, depending on the total emissions inventory characteristics (e.g., expected asymmetry), the probability distributions addressed in this study can be considered for the total uncertainty of the inventory, broadening the spectrum of possibilities already used.

Finally, it is considered that the present study provides the expected guidance for the interpretation and manipulation of emission factor standard uncertainties and will hopefully ease the process of implementing uncertainty estimation in GHG inventories and obtain a more precise and reliable quantification of emission data from the fossil-fueled transportation sector in Costa Rica and other countries around the world.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

REFERENCES

Abel, G. (2015). fanplot: an R package for visualizing sequential distributions. R J. 7, 15–23. doi: 10.32614/RJ-2015-002
Allen, K., and Padgett-Vásquez, S. (2017). Forest cover, development, and sustainability in Costa Rica: can one policy fit all? Land Use Policy 67, 212–221. doi: 10.1016/j.landusepol.2017.05.008
Audi, G., Kondev, F. G., Wang, M., Huang, W. J., and Naimi, S. (2017). The Nubase2016 evaluation of nuclear properties. Chin. Phys. C. 41:030001. doi: 10.1088/1674-1137/41/3/030001
Azzalini, A. (2020). The Skew-Normal and Related Distributions Such as the Skew-t [version 1.6-2]. Available online at: http://azzalini.stat.unipd.it/SN (Accessed 21, Oct 2020)
Azzalini, A., and Capitanio, A. (2014). The Skew-Normal and Related Families. Cambridge: Cambridge University Press.
Barlow, R. (2004). Asymmetrical Statistical Errors. Available online at: https://arxiv.org/abs/physics/0406120 (Accessed 21 Oct 2020)
Bharvirkar, R. (1999). Quantification of Variability and Uncertainty in Emission Factors and Emissions Inventories. Master’s thesis, North Carolina State University, Raleigh, NC.
Carnell, R. (2019). Triangle: Provides the Standard Distribution Functions for the Triangle Distribution [Version 0.12]. Available online at: https://bertcarnell.github.io/triangle/ (Accessed 17 Oct 2020)
Choula, M., Janssens-Maenhout, G., Super, L., Agusti-Panareda, A., Balsamo, G., Bousserez, N., et al. (2020). Global Anthropogenic CO2 Emissions and Uncertainties as Prior for Earth System Modelling and Data Assimilation. Available online at: https://doi.org/10.5194/essd-2020-68 (Accessed 11 Mar 2021).

AUTHOR CONTRIBUTIONS

GM-C contributed to the conception of the study, performed the statistical analysis, and wrote the first draft of the manuscript. BC-J revised it critically for important intellectual content and approved for publication of the content. Both authors contributed to the article and approved the submitted version.

FUNDING

The publication of this study was covered by the NDC Support Programme of the United Nations Development Programme (UNDP).

ACKNOWLEDGMENTS

We would like to thank Dr. Antonio Possolo (Statistical Engineering Division, National Institute of Standards and Technology, NIST) for his valuable guidance on the topic and general review of the manuscript. We also wish to thank Dr. Gilbert Brenes-Camacho (School of Statistics, Costa Rican University, UCR) for his technical review of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2021.662052/full#supplementary-material

Courvisanos, J., and Ameeta, J. (2006). A framework for sustainable Ecotourism: application to Costa Rica. Tour. Hosp. Plan. Dev. 3, 131–142. doi: 10.1080/14790530600938378
DCC and LCM (2020). Guía Metodológica para la Estimación y análisis de la Incertidumbre de Emisiones y Remociones de gases de Efecto Invernadero (GEI). National Carbon Neutrality PROGRAM 2.0. Available online at: https://cambioclimatico.go.cr/wp-content/uploads/2020/11/PPCN-Guiaincertidumbre.pdf (Accessed 23 Nov 2020)
DCC and PMR (2020). Programa País de Carbono Neutralidad: Categoría Organizacional. National Carbon Neutrality Program 2.0. Available online at: https://cambioclimatico.go.cr/wp-content/uploads/2020/04/1-PPCN_Organizacional.pdf (Accessed 10 Oct 2020)
Dennis, J. E., and Schnabel, R. B. (1996). Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Philadelphia, PA: SIAM.
EPA (1996). Evaluating the uncertainty of emission estimates. Final Report. Emission Inventory Improvement Program. Available online at: https://www.epa.gov/sites/production/files/2015-08/documents/vi04.pdf (Accessed 05 Oct 2020)
EPA (2002). Quality Assurance / Quality Control and Uncertainty Management Plan for the US Greenhouse Gas inventory: Procedures Manual for Quality Assurance / Quality Control and Uncertainty Analysis. Available online at: https://teipis.epa.gov/Exei/ZyPURL.cgi?DocKey=Pi005GXIH.TXT (Accessed 05 Oct 2020)
Fendler, L. (2017). All that Glitters is not Green: Costa Rica’s Renewables Conceal Dependence on Oil. Available online at: https://www.theguardian.com/world/2017/jan/05/costa-rica-renewable-energy-oil-cars (Accessed 26 Jan 2021)
Frey, C. (2003). Evaluation of an Approximate Analytical Procedure for Calculating Uncertainty in the Greenhouse Gas Version of the Multi-Scale Motor Vehicle and Equipment Emissions System. Ann Arbor, MI: U.S. Environmental Protection Agency.
Frey, C. (2007). *Quantification of Uncertainty in Emission Factors and Inventories*. Emission Inventory Conference, NC. Available online at: https://www3.epa.gov/tnichiel/conference/e16/session5/frey.pdf (Accessed 06 Oct 2020).

Government of Costa Rica (2019). *National Decarbonization Plan (Plan Nacional de Descarbonización)*. Available online at: https://cambioclimatico.go.cr/wp-content/uploads/2019/02/PLAN.pdf (Accessed 26 Jan 2021).

Hasselman, B. (2018). *nleqslv: Solve Systems of Nonlinear Equations [version 3.3.2]*. Available online at: https://cran.R-project.org/package=nleqslv (Accessed 20 Oct 2020).

ICE Group (2020). *We Are Renewable and Solidary Energy*. Available online at: https://www.grupoice.com/wps/wcm/connect/579df1f-5156-41e0-807d-d8f8085d718/Fasciculo_Electricidad_2020英格%C3%A9s_compressed.pdf?MOD=AIPERES&CVID=m.pGzcp (Accessed 26 Jan 2021).

IMN (2020). *Factores de Emisión de Gases de efecto Invernadero, 10th Edn.* National Meteorology Institute of Costa Rica. Available online at: http://cglobal.imn.ac.cr/index.php/publications/factores-de-emision-gei-decima-edicion-2020/ (Accessed 16 Oct 2020).

Intergovernmental Panel on Climate Change [IPCC] (2000). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories. Quantifying Uncertainties in Practice*. Intergovernmental Panel on Climate Change, Methodology Report for the National Greenhouse Gas Inventories Programme. Available online at: https://www.ipcc.ch/publication/good-practice-guidance-and-uncertainty-management-in-national-greenhouse-gas-inventories/ (Accessed 01 Oct 2020).

Intergovernmental Panel on Climate Change [IPCC] (2006). *Volume 1: General Guidance and Reporting – Uncertainties*. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Available online at: https://www.ipcc-nggip.iges.or.jp/public/2006gl/vol1.html (Accessed 01 Oct 2020).

ISO 14064-1. (2006). *Greenhouse Gases — Part 1: Specification with Guidance at the Organization Level for Quantification and Reporting of Greenhouse gas Emissions and Removals*. Geneva: International Organization for Standardization.

ISO 14064-1. (2018). *Greenhouse Gases — Part 1: Specification with Guidance at the Organization Level for Quantification and Reporting of Greenhouse Gas Emissions and Removals*. Geneva: International Organization for Standardization.

Joint Committee for Guides in Metrology [JCGM] (2008). *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement GUM: 1995 with minor corrections*. Joint Committee for Guides in Metrology, JCGM:100. Available online at: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf. (Accessed 01 Nov 2020).

Joint Committee for Guides in Metrology [JCGM] (2012). *International Vocabulary of Metrology – Basic and general concepts and associated terms (VIM): 2008 with minor corrections*. Joint Committee for Guides in Metrology, JCGM:200. Available online at: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf. (Accessed 08 Mar 2021).

Jonas, M., White, T., Marland, G., Lieberman, D., Nahorski, Z., and Nilsson, S. (2010). *Dealing with uncertainty in GHG inventories: how to go about it?* in *Coping with Uncertainty*. Lecture Notes in Economics and Mathematical Systems, Vol. 633, eds K. Marti, Y. Ermolieff, and M. Makowski 229–245. doi: 10.1007/978-3-7375-1_11

Milne, A., Glendining, M., Lark, R. M., Perryman, S., Gordon, T., and Whitmore, A. (2015). Communicating the uncertainty in estimated greenhouse gas emissions from agriculture. *J. Environ. Manag.* 160, 139–153. doi: 10.1016/j.jenvman.2015.05.034

Pettic, N. W., and Dye, S. (2013). *Notes on Triangular Distributions*. Statistics Learning Centre. Available online at: https://learnandteachstatistics.files.wordpress.com/2013/07/notes-on-triangle-distributions.pdf (Accessed 18 Oct 2020).

Posolos, A., and Iyer, H. (2017). Concepts and tools for the evaluation of measurement uncertainty. *Rev. Sci. Instrum.* 88:011301. doi: 10.1063/1.4974274

Posolos, A., Merkatas, C., and Bodnar, O. (2019). Asymmetrical uncertainties. *Metrologia* 56:045009. doi: 10.1088/1681-7575/ab2a6d

R Core Team (2020). *R: A language and Environment for Statistical Computing [version 3.6.3]*. Vienna: R Foundation for Statistical Computing.

Ritter, K., Lev-On, M., and Shires, T. (2010). Understanding uncertainty in greenhouse gas emission estimates: Technical considerations and statistical calculation methods. *Paper Presented at the 19th Annual International Emission Inventory Conference, Texas.*

Rypdal, K., and Winiewater, W. (2001). Uncertainties in greenhouse gas emission inventories – evaluation, comparability and implications. *Environ. Sci. Policy 4*, 107–116. doi: 10.1016/S1462-9011(00)00113-1

Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., and Janssens-Maenhout, G. (2020). *Uncertainties in the EDGAR Emission Inventory of Greenhouse Gases*. Available online at: https://doi.org/10.5194/acp-2020-1102 (Accessed 11 Mar 2021).

Stephenson, A. G. (2002). *evd: extreme value distributions*. R Core Team (2020). *R: A language and Environment for Statistical Computing [version 3.6.3]*. Vienna: R Foundation for Statistical Computing.

Wallis, K. (2014). The two-piece normal, bimodal, or double Gaussian distribution: its origin and rediscoveries. *Statist. Sci.* 29, 106–112. doi: 10.1214/13-STS417

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2021 Molina-Castro and Calderón-Jiménez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*