Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990-2019: Updates for Natural Gas and Petroleum Systems CO_2 Uncertainty Estimates

EPA updated the approach to estimate uncertainty for CH_4 emissions from natural gas and petroleum systems in the 2018 *Inventory of U.S. Greenhouse Gas Emissions and Sinks* (GHGI). EPA previously did not calculate the uncertainty for CO_2 emissions specifically, but instead applied the CH_4 uncertainty bounds to the estimated CO_2 emissions. This memorandum discusses an update included in the 2021 GHGI to calculate uncertainty bounds specific to CO_2 emissions from National Gas and Petroleum Systems.

1 Background and 2020 (Previous) GHGI Methodology

For each annual GHGI, EPA conducts a quantitative uncertainty analysis using IPCC Approach 2 methodology (i.e., Monte Carlo simulations technique). IPCC suggests the use of a 95% confidence interval, which is the interval that has a 95% probability of containing the unknown "true" value. Therefore, EPA uses @RISK, a Microsoft Excel add-in tool to estimate the 95% confidence bound around CH_4 emissions from both the natural gas and petroleum systems inventories. Due to the significant number of emissions sources in natural gas and petroleum systems (i.e., each contains more than 100 emission sources), EPA does not calculate the uncertainty for every emission source. Rather, EPA calculates the uncertainty for the highest-emitting sources that cumulatively contribute at least 75% of gross emissions in natural gas and petroleum systems in the most recent GHGI year, and then applies those results via Monte Carlo simulations to the emissions for the other smaller sources to estimate the overall uncertainty. The 75% cumulative contribution was determined, through the stakeholder process, to be an appropriate level of precision given the large number of emission sources included in both the natural gas systems and petroleum systems.

In previous GHGIs, prior to 2021, EPA did not calculate uncertainty bounds specific to CO_2 emissions. Instead, EPA applied the calculated CH_4 bounds for natural gas and petroleum systems inventories, expressed as the percent (%) deviation above and below, to the CO_2 emissions estimates.

To develop a 95% confidence interval for an emission estimate from a chosen sector (e.g., natural gas systems), it is necessary to characterize the probability density function (PDF) of the average emission and activity factors for each emission source contributing to that source category emission estimate. The PDF describes the range and relative likelihood of possible values for the average emission and activity factors corresponding to that emission source (e.g., flares in the natural gas processing segment). EPA develops uncertainty model parameters based on published studies, Greenhouse Gas Reporting Program (GHGRP) Subpart W data, and/or expert judgment for each of the top emission sources. If the modeling input (e.g., emission factor) is based on GHGRP Subpart W data, EPA employs bootstrapping to determine the shape and other parameters of the sampling distribution of the mean value. The bootstrapping analysis enables the determination of the PDF (e.g., normal, lognormal) as well as applicable statistical parameters (e.g., standard deviation, maximum, minimum) needed for the Monte Carlo simulation. For modeling inputs based on recently published studies (e.g., Zimmerle et al. 2019), EPA directly uses uncertainty information included in the study.[1] For modeling inputs based on older data sets (e.g., 1996 EPA/GRI study) or macro parameters, which are used as inputs to several emission source estimates (e.g., total active well counts from Enverus DrillingInfo), EPA treats these input parameters as a uniformly distributed estimate and refers to published estimates and expert judgment to estimate upper and lower bounds. For input values obtained from certain data sources where uncertainty data are not available, EPA assigns uncertainty bounds based on expert

[1] Gathering and boosting CH_4 emissions were a top source in the 2020 GHGI uncertainty analyses. Zimmerle, Daniel et al., Characterization of Methane Emissions from Gathering Compressor Stations. Available at https://mountainscholar.org/handle/10217/195489. October 2019.

judgment based on a characterized level of confidence; for example, EPA assigns uncertainty bounds of 5% to the U.S. Energy Information Administration (EIA) data.

Per the Intergovernmental Panel on Climate Change (IPCC) Guidance, an uncertainty analysis should be seen as a means to help prioritize national efforts to reduce the uncertainty of inventories in the future, and guide decisions on methodological choice.[2] Uncertainty estimates in the GHGI capture quantifiable uncertainties in the input activity and emission factors data, but do not account for the potential of additional sources of uncertainty such as modeling uncertainties, data representativeness, measurement errors, and misreporting or misclassification.

2 CO_2 Uncertainty Analysis

EPA updated the uncertainty methodology for the 2021 GHGI and applied the Monte Carlo simulation technique to calculate the 95% confidence interval for CO_2 emissions in natural gas and petroleum systems. For this initial CO_2 uncertainty analysis, EPA examined year 2018 emissions from the 2020 (previous) GHGI and did not update the analysis to use year 2019 emissions from the 2021 GHGI. The CO_2 uncertainty bounds (expressed as a percent) calculated for year 2018 in the 2020 GHGI were applied to year 2019 emissions in the 2021 GHGI, as shown in Table 8.

As a first step, EPA reviewed the 2020 (previous) GHGI CO_2 emissions for year 2018 to assess the highest-emitting sources and identify those that cumulatively contribute at least 75% of emissions. Table 1 and Table 2 show the top 15 sources of 2018 emissions for natural gas and petroleum systems, respectively.

Table 1. Top 15 Sources of CO_2 Emissions for Natural Gas Systems in 2020 (Previous) GHGI

Industry Segment	Emission Source	2018 CO_2 Emissions (mt)	% of Total CO_2 Emissions	% of Total CO_2 Emissions, Cumulative	Source in top 75%?
Processing	Acid Gas Removal (AGR) Vents	17,451,105	49.9%	49.9%	Yes
Processing	Flares	6,981,114	20.0%	69.9%	Yes
Production	G&B Stations - Flare Stacks	4,205,760	12.0%	81.9%	Yes
Production	Miscellaneous Onshore Production Flaring	1,380,268	3.9%	85.8%	
Production	G&B Stations - Tanks	1,294,821	3.7%	89.5%	
Production	Condensate Tanks	844,923	2.4%	92.0%	
Production	G&B Stations - Dehydrators	801,603	2.3%	94.2%	
Production	G&B Stations - AGR	643,969	1.8%	96.1%	
Exploration	HF Completions	391,897	1.1%	97.2%	
LNG Export	LNG Export Terminals	273,956	0.8%	98.0%	
Production	Pneumatic Controllers	111,831	0.3%	98.3%	
Production	HF Workovers	106,196	0.3%	98.6%	
Transmission + Storage	Flaring (Storage)	80,016	0.2%	98.8%	
Transmission + Storage	Flaring (Transmission)	75,251	0.2%	99.1%	
Production	G&B Stations - other	70,463	0.2%	99.3%	
TOTAL		**34,971,601**			

[2] 2006 IPCC Guidelines for National Greenhouse Gas Inventories; Chapter 3 - Uncertainties. https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/1_Volume1/V1_3_Ch3_Uncertainties.pdf

Table 2. Top 15 Sources of CO_2 Emissions for Petroleum Systems in 2020 (Previous) GHGI

Industry Segment	Emission Source	2018 CO_2 Emissions (mt)	% of Total CO_2 Emissions	% of Total CO_2 Emissions, Cumulative	Source in top 75%?
Production	Associated Gas Flaring	18,980,470	51.6%	51.6%	Yes
Production	Oil Tanks	6,369,067	17.3%	68.9%	Yes
Production	Miscellaneous Production Flaring	4,226,320	11.5%	80.3%	Yes
Refinery	Flaring	3,648,222	9.9%	90.2%	
Exploration	HF Well Completions	2,729,682	7.4%	97.7%	
Production	Offshore Facilities (GoM Federal)	411,412	1.1%	98.8%	
Production	Offshore Facilities (AK)	122,362	0.3%	99.1%	
Production	HF Workovers	92,895	0.3%	99.4%	
Production	Pneumatic Controllers	81,375	0.2%	99.6%	
Refinery	Process Vents	53,693	0.1%	99.7%	
Refinery	Asphalt blowing	32,559	0.1%	99.8%	
Exploration	Non-completion Well Testing	31,698	0.1%	99.9%	
Production	Offshore Facilities (Pacific)	8,688	<0.05%	99.9%	
Production	Chemical Injection Pumps	7,834	<0.05%	100.0%[a]	
Production	Associated Gas Venting	5,484	<0.05%	100.0%[a]	
TOTAL		**36,814,372**			

a. Cumulative emissions are less than 100%, but value is rounded to show to one decimal point.

Flaring and acid gas removal (AGR) emissions are the primary source of CO_2 emissions in natural gas and petroleum systems and most of the top individual emission sources include either a flare or an AGR unit. Based on year 2018 emissions, each sector has one emission source that accounts for approximately 50% of total CO_2 emissions: processing plant AGR units for natural gas systems and associated gas flaring for petroleum systems. Each sector also needs only three emission sources to achieve the 75% emissions threshold for the uncertainty analysis. In general, the largest CO_2 emission sources are different than the largest CH_4 emission sources.

It should be noted that each of the flaring and AGR emission sources that cumulatively contribute at least 75% of emissions to natural gas and petroleum systems rely on emission factors and activity factors calculated from Subpart W data. In each of these instances, EPA used a bootstrapping analysis to characterize the PDF (e.g., normal, lognormal) and statistical parameters (e.g., standard deviation) for the Monte Carlo simulation. Bootstrapping analyses are further discussed in the following section. The uncertainty results from the sources that cumulatively contribute at least 75% of emissions were used to estimate the uncertainty for the other smaller emission sources and the overall uncertainty via Monte Carlo simulation (as discussed in Section 1).

2.1 Bootstrapping Results

EPA performed the bootstrapping analyses for each of the Subpart W emission factors (EFs) and activity factors (AFs). Results are provided below. Table 3 and Table 4 provide the GHGI mean value for the year 2018, the PDF and relevant inputs for the Monte Carlo simulation as determined by the Microsoft Excel @RISK add-in tool, and the simulated 95% interval for the GHGI mean emission source EFs and AFs for natural gas (Table 3) and petroleum (Table 4) systems. The 95% interval is shown as the percent above and below the GHGI mean and is for contextual purposes only.

The PDF for each EF and AF was chosen using a best fit analysis performed in @RISK. This approach is slightly different than the current approach for natural gas and petroleum system CH_4 emissions, as well as the overall US GHGI uncertainty analysis, which both limit the possible PDF shapes to the most common types (e.g., normal, lognormal, etc.). The IPCC Guidance[3] notes that there can be large differences between different distribution functions at the extremes, where there are few or no data to constrain distribution type. This highlights the importance of identifying the PDF of best fit during this step of the uncertainty analysis. As the GHGRP data evaluated here are considered to be robust and large datasets, the EPA did not limit the PDF shapes fit by @RISK. EPA sought stakeholder feedback on this approach, but none was received. Table 5 shows an example of each PDF assigned by @RISK using a best fit function, a pictorial representation of that assigned shape, and a histogram with 1,000 datapoints as a result of the bootstrapping.

[3] 2006 IPCC Guidelines provide 'Good Practice Guidance' for selecting PDFs (Section 3.2.2.4). "In many cases, several functions will fit the data satisfactorily within a given probability limit. These different functions can have radically different distributions at the extremes where there are few or no data to constrain them, and the choice of one function over another can systematically change the outcome of an uncertainty analysis. Cullen and Frey (1999) reiterate the advice of previous authors in these cases that it must be knowledge of the underlying physical processes that governs the choice of a probability function. What the tests provide, in the light of this physical knowledge, is guidance on whether this function does or does not satisfactorily fit the data" (pg 24).

Table 3. Overview of Natural Gas Systems Year 2018 CO_2 Uncertainty Inputs for @RISK Modeling

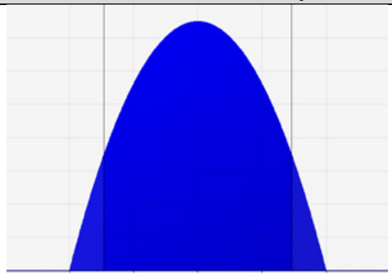
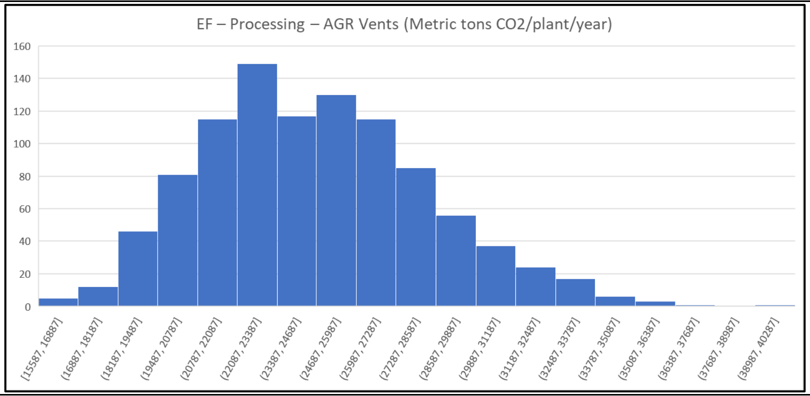
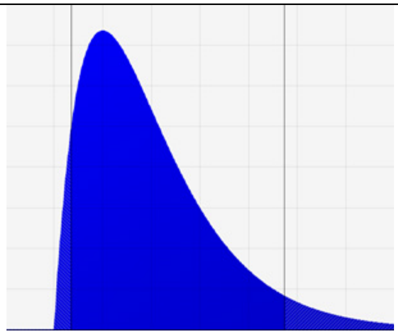
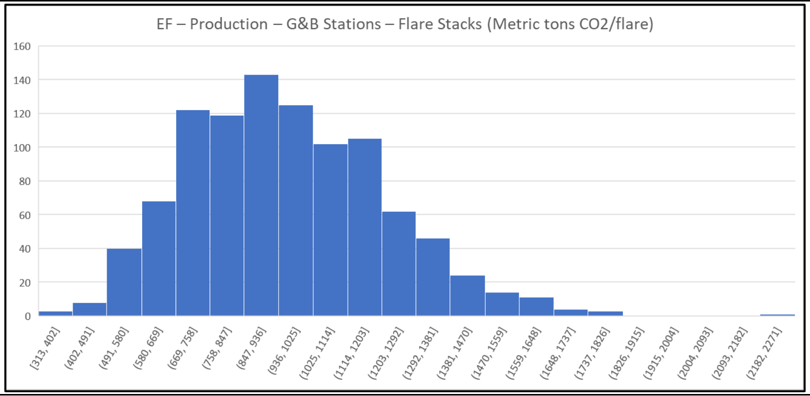
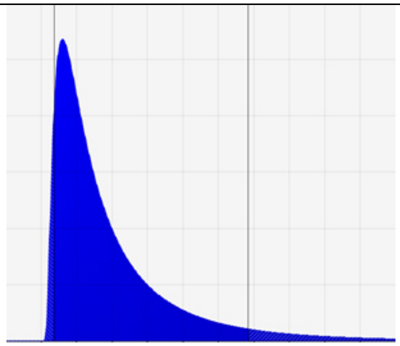
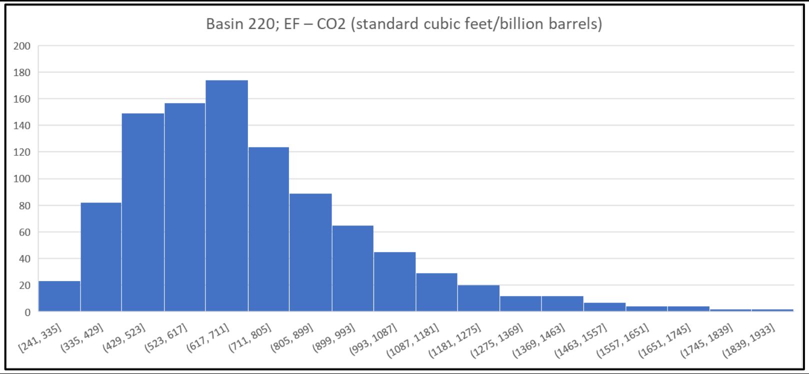
Emissions Calculation Input	Year 2018 GHGI Mean Value	PDF	Relevant Inputs	2.5% Percentile	97.5% Percentile
EF – Processing – AGR Vents (Metric tons CO2/plant/year)	24,771	Beta General	Shape Parameter 1 = 6.4 Shape Parameter 2 = 20 Min = 13,766 Max =58,076	18,565 (-25%)	32,572 (32%)
EF – Processing – Flares (Metric tons CO2/plant/year) [a]	10,466	Lognorm	Standard Deviation = 1,538 Shift = 1,179	7,752 (-26%)	13,831 (32%)
EF – Production – G&B Stations – Flare Stacks (Metric tons CO2/flare)	920	Gamma	Shape = 14 Scale = 67 Shift = -8.6	531 (-45%)	1,527 (59%)
AF – Production – G&B Stations – Flare Stacks (flare count)	4,254	Gamma	Shape = 6.5 Scale = 341 Shift = 1,992	2,839 (-33%)	6,215 (47%)

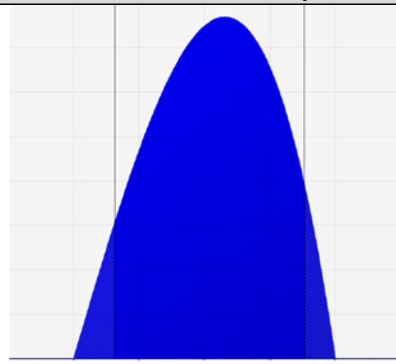
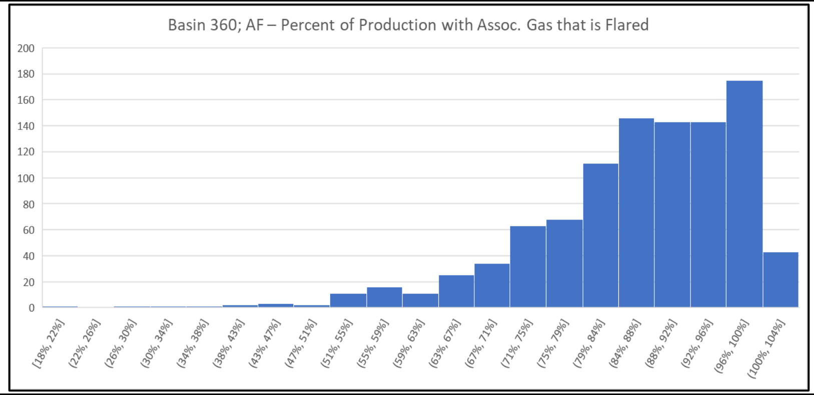
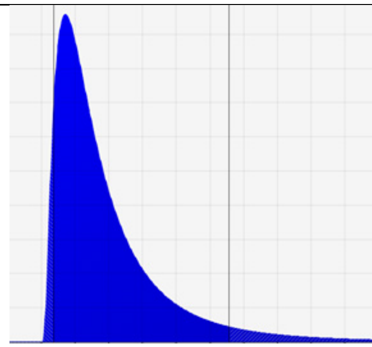
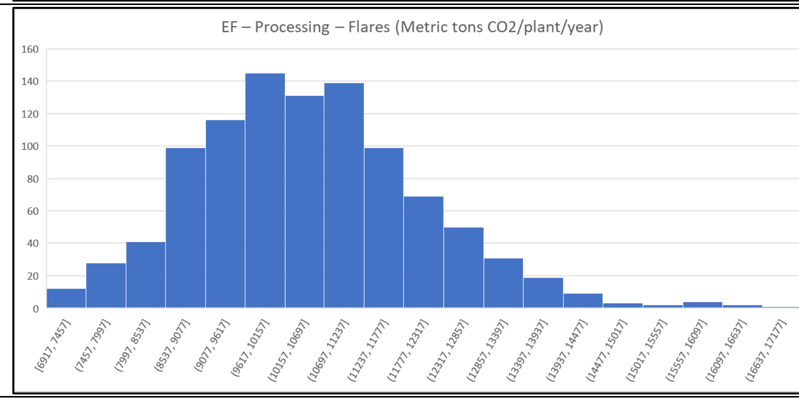
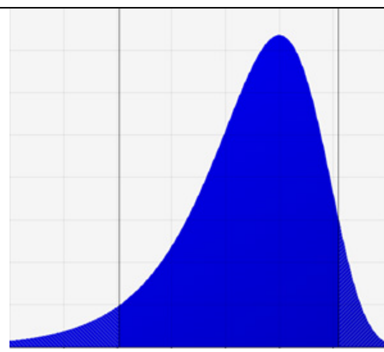
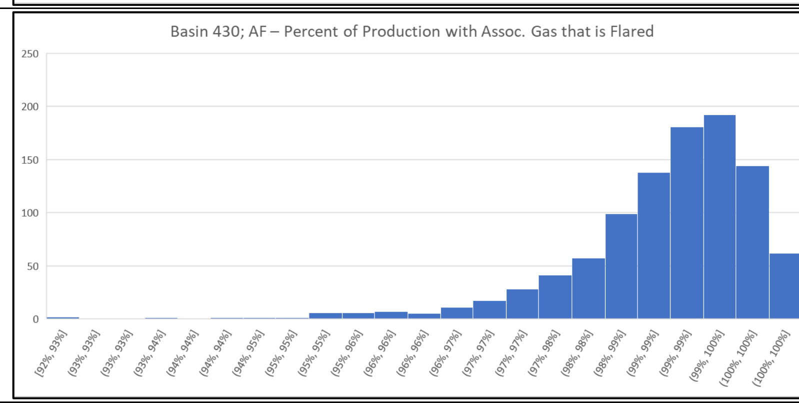
Table 4. Overview of Petroleum Systems Year 2018 CO_2 Uncertainty Inputs for @RISK Modeling

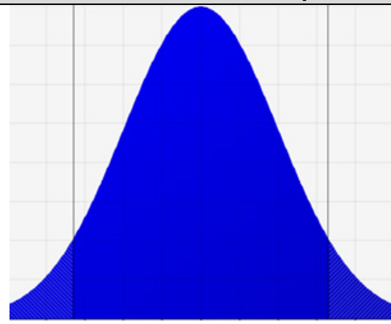
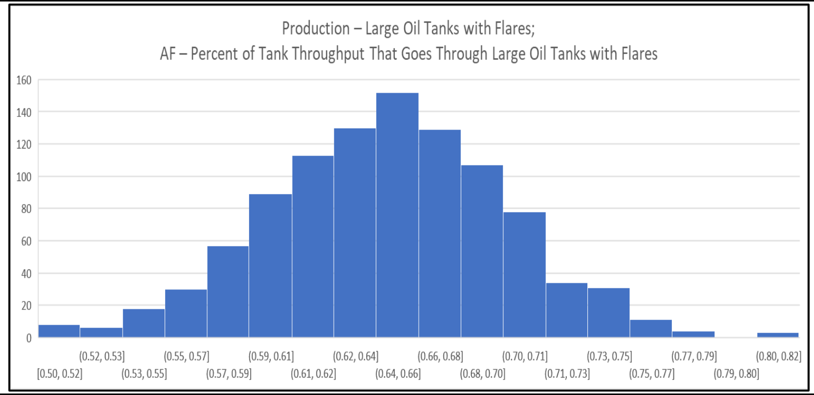
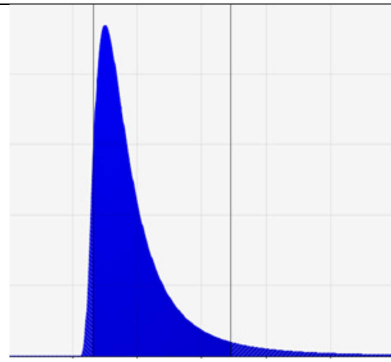
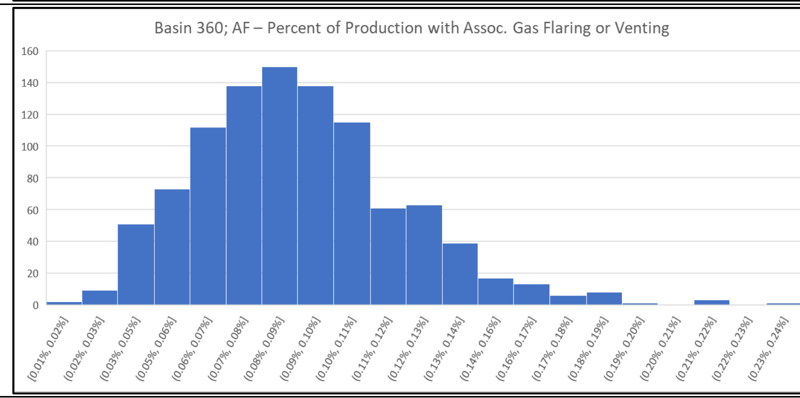
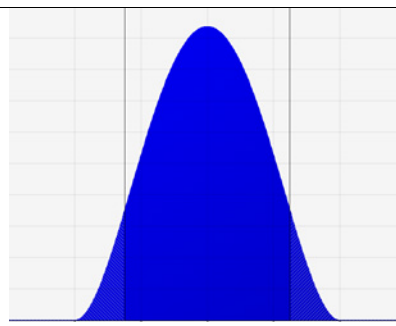
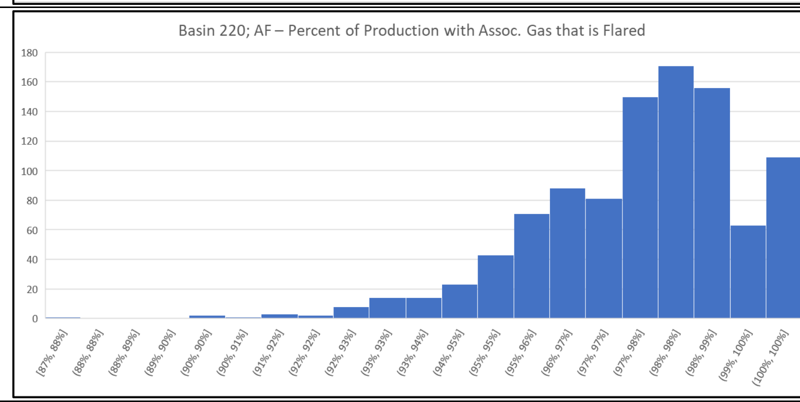
Emissions Calculation Input		Year 2018 GHGI Mean Value	PDF	Relevant Inputs	2.5% Percentile	97.5% Percentile
Production – Associated Gas Flaring						
Basin 220	AF – Percent of Production with Assoc. Gas Flaring or Venting	3.9%	Lognorm	Mean = 0.047 Standard Deviation = 0.022 Shift = -0.0050	1.3% (-69%)	9.5% (128%)
	AF – Percent of Production with Assoc. Gas that is Flared	97.6%	Pert	Min = 0.86 Most Likely Value for Shape = 1.0 Max = 1	93.1% (-5%)	99.9% (2%)
	EF – CO2 (standard cubic feet/billion barrels)	633	Invgauss	Mean = 653 Shape = 3,863 Shift = 60	340 (-52%)	1,423 (99%)
Basin 360	AF – Percent of Production with Assoc. Gas Flaring or Venting	0.09%	Pearson5	Shape: 36 Scale: 0.065 Shift: -0.00094	0.03% (-60%)	0.15% (79%)
	AF – Percent of Production with Assoc. Gas that is Flared	86.5%	Kumaraswamy	Shape Parameter 1 = 1.8 Shape Parameter 2 = 0.33 Min = 0.18 Max = 1.0	55.3% (-36%)	100% (17%)
	EF – CO2 (standard cubic feet/billion barrels)	5,987	Gamma	Shape = 7.7 Scale = 1,016 Shift = -1,798	1,492 (-75%)	11,899 (97%)
Basin 395	AF – Percent of Production with Assoc. Gas Flaring or Venting	58.8%	Gamma	Shape = 45 Scale = 0.016 Shift = -0.12	41.3% (-32%)	83.2% (38%)
	AF – Percent of Production with Assoc. Gas that is Flared	100%	Kumaraswamy	Shape Parameter 1 = 1.0 Shape Parameter 2 = 0.20 Min = 1.0 Max = 1.0	100% (-0.02%)	100% (0.01%)
	EF – CO2 (standard cubic feet/billion barrels)	683	Beta General	Shape Parameter 1 = 4.6 Shape Parameter 2 = 16 Min = 331 Max = 1,960	453 (-34%)	1,007 (46%)
Basin 430	AF – Percent of Production with Assoc. Gas Flaring or Venting	37.8%	Weibull	Shape = 2.0 Scale = 0.36 Shift = 0.065	13.1% (-66%)	76.8 (99%)
	AF – Percent of Production with Assoc. Gas that is Flared	99.0%	Minimum Extreme Value	Location = 0.99 Shape = 0.0065	96.1% (-3%)	99.9% (1%)
	EF – CO2 (standard cubic feet/billion barrels)	293	Invgauss	Mean = 327 Shape = 1,185 Shift = 20	130 (-62%)	769 (121%)

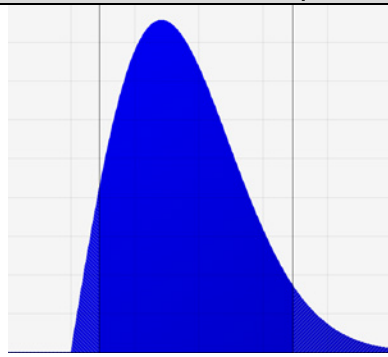
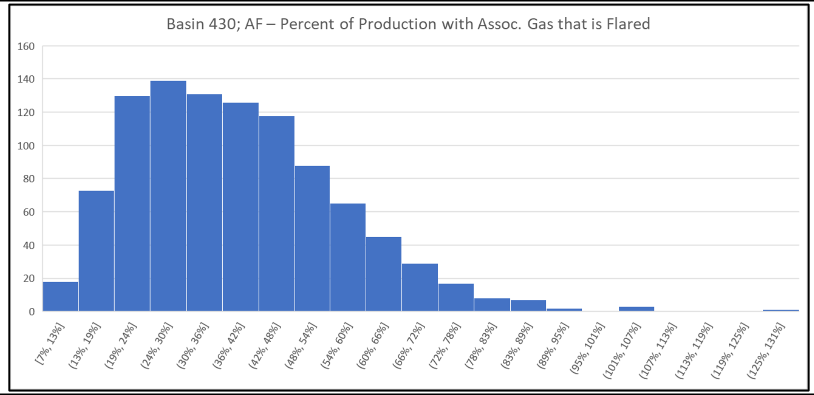
Emissions Calculation Input		Year 2018 GHGI Mean Value	PDF	Relevant Inputs	2.5% Percentile	97.5% Percentile
Other Basins	AF – Percent of Production with Assoc. Gas Flaring or Venting	4.2%	Gamma	Shape = 4.3 Scale = 0.0089 Shift = 0.0053	1.7% (-61%)	8.8% (102%)
	AF – Percent of Production with Assoc. Gas that is Flared	92.5%	Pert	Min = 0.52 Most Likely Value for Shape = 1.0 Max = 1.0	74.5% (-19%)	99.9% (9%)
	EF – CO2 (standard cubic feet/billion barrels)	450	Weibull	Shape = 2.0 Scale = 446 Shift = 108	185 (-63%)	956 (90%)
Production – Large Oil Tanks with Flares						
AF – Percent of Tank Throughput That Goes Through Large Oil Tanks with Flares		64.7%	Normal	Mean = 0.65 Standard Deviation = 0.050	54% (-16%)	75% (15%)
EF – CO2 (standard cubic feet/billion barrels)		87.4	Gamma	Shape = 24 Scale = 3.0 Shift = 16	62 (-30%)	119 (35%)
Miscellaneous Production Flaring						
Basin 220	EF – CO2 (Metric tons/billion barrels)	0.0011	Pearson5	Shape = 33 Scale = 0.070 Shift = -0.0010	0.0005 (-54%)	0.002 (76%)
Basin 395	EF – CO2 (Metric tons/billion barrels)	0.0035	Pert	Min = 0.000027 Most Likely Value for Shape = 0.000027 Max = 0.020	0.0001 (-96%)	0.0101 (194%)
Basin 430	EF – CO2 (Metric tons/billion barrels)	0.0009	Gamma	Shape = 19 Scale = 0.000078 Shift = -0.00051	0.0004 (-61%)	0.0017 (76%)
Other Basins	EF – CO2 (Metric tons/billion barrels)	0.0007	Invgauss	Mean = 0.0010 Shape = 0.023 Shift = -0.00033	0.0003 (-50%)	0.0012 (70%)

Table 5. PDF Supplemental Information

PFD Types	PDF Pictorial Representation	Example, Emission Calculation Input PDF
Beta General		
Gamma		
Invgauss		

PFD Types	PDF Pictorial Representation	Example, Emission Calculation Input PDF
Kumaraswamy		
Lognorm		
Minimum Extreme Value		

PFD Types	PDF Pictorial Representation	Example, Emission Calculation Input PDF
Normal		
Pearson5		
Pert		

PFD Types	PDF Pictorial Representation	Example, Emission Calculation Input PDF
Weibull		

2.2 Monte Carlo Results

Tables 6 and 7 summarize the calculated source category level uncertainty estimates for petroleum and natural gas systems based on year 2018 CO_2 emissions from the 2020 (previous) GHGI. Included as the last row in each table is the methane uncertainty results from last year's GHGI for comparison. These Monte Carlo results based on year 2018 CO_2 emissions in the 2020 GHGI were applied to year 2019 emissions in the 2021 GHGI; see Table 8. In future GHGIs the uncertainty estimates will be quantified for the most recent year of data.

Table 6. Summary of Petroleum Systems Year 2018 CO_2 Uncertainty Results

Emission Source		Mean Year 2018 Emissions (MT CO_2)	2.5% Lower Bound of Mean Year 2018 Emissions (MT CO_2)		97.5% Upper Bound of Mean Year 2018 Emissions (MT CO_2)	
			Value	%	Value	%
Associated Gas Flaring	220 Gulf Coast	686,281	162,148	-76%	1,859,022	171%
	360 Anadarko	37,482	6,334	-83%	100,827	169%
	395 Williston	10,131,704	5,636,295	-44%	16,568,499	64%
	430 Permian	7,248,710	1,548,479	-79%	20,402,926	181%
	Other	876,292	204,927	-77%	2,237,256	155%
Production – Large Oil Tanks with Flares		6,369,067	4,315,997	-32%	8,963,286	41%
Miscellaneous Production Flaring	220 Gulf Coast	686,842	305,011	-56%	1,223,911	78%
	395 Williston	1,653,170	62,280	-96%	5,152,768	212%
	430 Permian	1,182,863	455,995	-61%	2,086,585	76%
	Other	703,446	338,856	-52%	1,199,926	71%
Total for Sources Modeled [a]		29,575,857	20,514,329	-31%	43,877,159	48%
Total for Sources Not Modeled		7,238,515	4,643,803	-36%	10,697,527	48%
Source Category Total		**36,814,372**	**26,890,336**	**-27%**	**51,923,681**	**41%**

a. Those sources that cumulatively contribute at least 75% of emissions.

Table 7. Summary of Natural Gas Systems Year 2018 CO_2 Uncertainty Results

Emission Source	Mean Year 2018 Emissions (MT CO_2)	2.5% Lower Bound of Mean Year 2018 Emissions (MT CO_2)		97.5% Upper Bound of Mean Year 2018 Emissions (MT CO_2)	
		Value	%	Value	%
Acid Gas Removal Vents	16,522,287	12,304,773	-26%	21,834,132	32%
Flares	6,981,114	5,218,862	-25%	9,211,674	32%
Gathering & Boosting – Flare Stacks	4,205,760	1,997,940	-52%	7,567,846	80%
Total for Sources Modeled [a]	27,709,161	22,338,840	**-19%**	**34,076,332**	**23%**
Total for Sources Not Modeled	7,262,440	**5,675,255**	**-22%**	**8,865,628**	**22%**
Source Category Total	**34,971,601**	**29,295,317**	**-16%**	**41,463,998**	**19%**

a. Those sources that cumulatively contribute at least 75% of emissions.

Table 8. Summary of 2021 GHGI CO_2 Uncertainty Results

Sector	Mean Year 2019 Emissions (MMT CO_2)	2.5% Lower Bound of Mean Year 2019 Emissions (MMT CO_2)		97.5% Upper Bound of Mean Year 2019 Emissions (MMT CO_2)	
		Value	%	Value	%
Petroleum Systems	47.3	34.5	-27%	66.6	+41%
Natural Gas Systems	37.2	31.3	-16%	44.3	+19%

3 Requests for Stakeholder Feedback

EPA sought stakeholder feedback in the November 2020 memo and in the public review draft of the GHGI, but did not receive any stakeholder comments.

The questions below were not updated for this memorandum and are copied from the November 2020 memo.

Questions to Stakeholders

EPA seeks stakeholder feedback on the approach under consideration and the questions below.

1. EPA seeks general feedback on the approach of calculating uncertainty bounds for CO_2 emissions separately from CH_4 emissions.
2. EPA seeks feedback on applying the CH_4 emissions uncertainty methodology to CO_2 emissions (e.g., calculate the uncertainty for the highest-emitting sources that cumulatively account for at least 75% of total CO_2 emissions and use Monte Carlo simulations to calculate the uncertainty for the other smaller sources and the overall uncertainty).
3. EPA seeks feedback on whether the PDFs incorporated into the uncertainty analysis should be limited (e.g., normal, lognormal, uniform, triangular, and beta) or if other distributions should be considered (e.g., Weibull, Kumaraswamy, Pearson5).