Forestry and Land Use Change sector: Emissions from Water Reservoirs

Nicole Brown[1,3], Gabriella Volpato[2,3], and Elizabeth Reilly[1,3]

1) The Johns Hopkins University Applied Physics Laboratory,
2) WattTime, 3) Climate TRACE

1. Introduction

Water reservoirs play an important role in supporting human populations, providing hydroelectric power and supplying water for drinking and irrigation. However, they have also been large sources of anthropogenic greenhouse gas (GHG) emissions, emitting carbon and methane gasses generated by the breakdown of organic material in water. In 2019, the Intergovernmental Panel on Climate Change (IPCC) divided reservoirs into two categories of wetlands (Buendia *et al.*, 2019):

a. Land converted to flooded land (LCFL): reservoirs no more than 20 years old that were formed by flooding dry land or by expanding an existing smaller waterbody
b. Flooded land remaining flooded (FLRF): reservoirs that have been flooded for at least 20 years

Throughout literature, there have been some differences in the exact definition of reservoir. Frequently, the term has been used interchangeably to refer to both completely human-made water bodies that were previously unflooded and to natural water bodies with dams, dikes, or other forms of artificial water regulation. For the calculation of anthropogenic emissions, the IPCC has only considered the subset of reservoirs "excluding areas that were unmanaged water bodies (lakes and rivers) or unmanaged wetlands prior to flooding" (Buendia *et al.*, 2019). Following this definition, only completely human-made water bodies were considered reservoirs for these emissions calculations. As such, this influenced how countries report reservoir emissions, which may not capture all the emissions stemming from LCFL and FLRF.

Reservoir emissions are the sum of carbon dioxide (CO_2) gases released by the breakdown of submerged organic matter and methane (CH_4) gases escaping from sediment and aquatic plants. Recently, there have been several advancements in the methods used to estimate reservoir emissions where direct measurements cannot easily be made. Many studies have focused on quantifying correlations between geographic and environmental variables and greenhouse gas fluxes in order to create equations to estimate emissions. The work of Deemer *et al.* (2016), for instance, examined variables such as reservoir age and chlorophyll concentrations and their relationship to directly measured fluxes. Additionally, Prairie *et al.* (2021)developed the G-res

model, a multiple linear regression model that considers several more reservoir attributes, in order to generate per-reservoir emissions estimates. However, the importance and correlation of each variable with emissions are currently open research topics. As such, emission estimates from previous studies have varied by over a factor of ten, between 0.32 to 6.6 billion metric tonnes CO_2 equivalent (CO_2eq) (Harrison *et al.*, 2021). Because methodology and resulting reservoir emissions estimates have had such a large variance, the Climate TRACE coalition has developed a methodology that follows IPCC's guidelines to create a robust approach to estimate reservoir emissions globally.

2. Materials and Methods

The Climate TRACE coalition estimated CO_2, CH_4, and CO_2eq emissions for years 2015 to 2022 for reservoirs from the Global Reservoir and Dam (GRanD) dataset (Lehner *et al.*, 2011) using the IPCC's 2019 refinement guidelines for tier 1 emissions from wetlands (Buendia *et al.*, 2019). For each reservoir, emissions were estimated by multiplying a climate zone-specific emissions factor (EF) by the reservoir's surface area for each reservoir within a country, then the estimates were aggregated together to calculate the country's total emissions. While N_2O is also emitted by water reservoirs, the IPCC has assigned these emissions to other sectors to avoid double counting. With the approach described here, N_2O was not reported.

2.1 Datasets employed

Generating emissions required aggregating data across several datasets: reservoir attributes were taken from HydroLAKES v1.0 (https://www.hydrosheds.org/products/hydrolakes; Messager *et al.*, 2016) and GRanD v1.3 (https://www.globaldamwatch.org/grand/; Lehner *et al.*, 2011), climate zone and weather information were taken from the Climatic Research Unit gridded Time Series v4.07 (CRU TS) (https://crudata.uea.ac.uk/cru/data/hrg/; Harris *et al.*, 2020), EFs were taken from the IPCC (Buendia *et al.*, 2019), and elevation data were taken from the Tropospheric Emission Monitoring Internet Service (TEMIS) 0.5-degree version of GMTED2010 (https://www.temis.nl/data/gmted2010.html; Geffen, 2023).

HydroLAKES is a global dataset of all lakes with a surface area of at least 0.1 km^2 (Messager *et al.*, 2016). GRanD is included within HydroLAKES and consists of 7,320 dams with reservoirs more than 0.1 km^3 in volume (Lehner *et al.*, 2011). In 2019, it was estimated to represent 75 percent of the total surface area of all reservoirs and was one of the few datasets that contained georeferenced reservoirs with detailed attribute information (Buendia *et al.*, 2019). HydroLAKES has labeled each waterbody in its dataset as a "lake", "reservoir", or a "lake control/natural lake with regulation structure". Only reservoirs in GRanD that were part of the lake control category were included for anthropogenic emissions calculations. Reservoirs that had no estimated completion year, or that were listed as under construction, destroyed, or removed as of 2022 were also omitted from our dataset, leaving 6,275 reservoirs with a total

surface area of 243,078 km^2. GRanD's "Area_skm" column was used as the surface area (i.e., capacity) for each reservoir, and the reservoir's type was taken from the "Main_use" column.

Emissions factors for reservoirs from the IPCC were provided separately for LCFL and FLRF in six different climate zones, aggregated from the 12 original zones listed in Table 1 (Buendia *et al.*, 2019). Climate zones were assigned to each reservoir based on temperature (TMP), frost day frequency (FRS), potential evapotranspiration (PET), and precipitation (PRE) from CRU TS in conjunction with elevation from GMTED2010. GMTED2010 is a global elevation model that combines data from various radar and satellite sources. The CRU TS is a global dataset of daily weather-related data from 1901 to 2022, and it was used by the IPCC to create their global map of climate zones (Buendia *et al.*, 2019).

Table 1. IPCC climate zones and their aggregations. The "j" column represents the numerical identifier for each aggregated climate zone type. The "n" column denotes the number of reservoirs in our dataset for each climate zone.

IPCC Climate Zones	Aggregated Climate Zones	j	n
Boreal dry	Boreal	1	59
Boreal moist			
Polar dry			
Polar moist			
Cool temperate dry	Cool temperate	2	1907
Cool temperate moist			
Warm temperate dry	Warm temperate dry	3	1700
Warm temperate moist	Warm temperate moist	4	1375
Tropical dry	Tropical dry/montane	5	541
Tropical montane			
Tropical moist	Tropical moist/wet	6	693
Tropical wet			

2.2 Methods

To determine the set of emissions factors to use, each reservoir was first mapped to its climate zone using its latitude and longitude found in GRanD. To assign this climate zone, we created an updated version of the IPCC's global mapping using the decision tree shown in the 1st

Corrigenda to the 2019 guidelines (Federici, 2021). While the IPCC used CRU TS data averaged across 1985 to 2015 to derive climate zones, our version used more recent data from 1993 to 2022. The decision tree required estimates for FRS, elevation, mean monthly temperature (MMT), mean annual temperature (MAT), mean annual precipitation (MAP), and mean annual precipitation to potential evapotranspiration ratio (MAP:PET). MMT, which is the same as TMP, and elevation were taken directly from CRU TS and GMTED2010, respectively. MAT and MAP were derived from 30-year averages of TMP and PRE, respectively. MAP:PET was simply calculated as MAP divided by PET. Because this data was all available at 0.5-degree resolution, the resulting map was also generated at this resolution, shown in Figure 1. Although climate remains relatively consistent over long periods of time, there were some notable differences over the IPCC version of the map, such as larger polar moist regions in Alaska and more tropical dry regions across the Sahara. The number of reservoirs in our dataset in each climate zone is shown in Table 1.

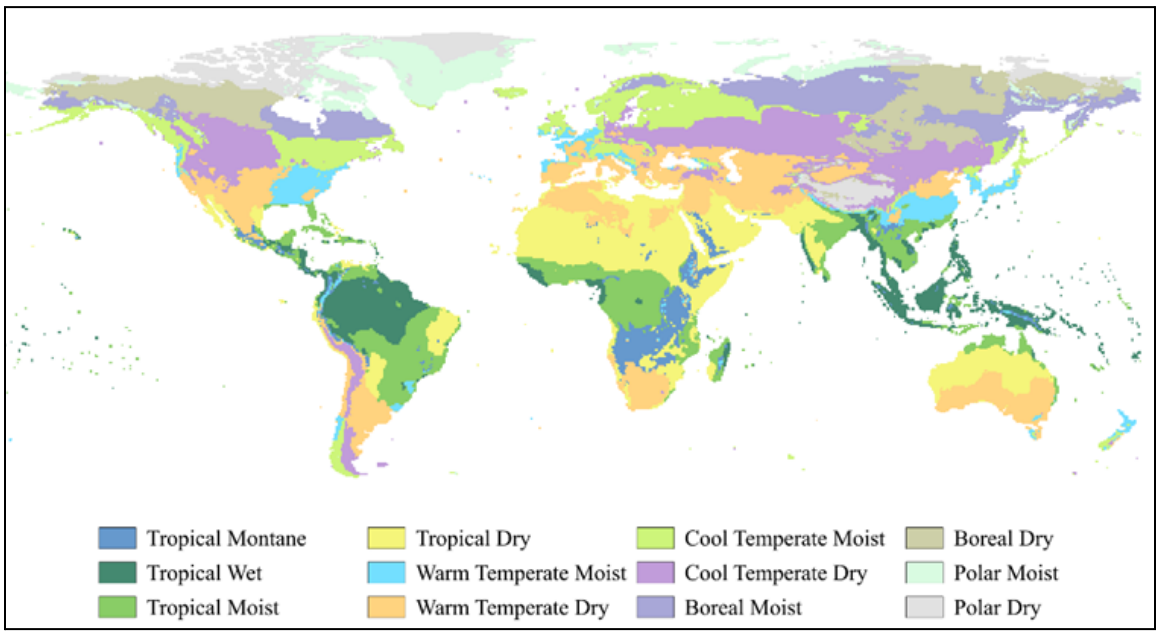

Figure 1. Recreated map of IPCC climate zones, with climate data from 1993-2022.

After being assigned a climate zone, reservoirs were labeled as FLRF or LCFL for each year from 2015 to 2022. All reservoirs with a completion date (i.e., the "Year" column in GRanD) no more than 20 years prior to the year of the emissions estimate were considered LCFL. All other reservoirs greater than 20 years old at the time of the estimate were considered FLRF. Note that FLRF and LCFL labels are year-specific. For instance, a reservoir built in 1996 would have been considered LCFL for 2015 emissions estimates and would have used LCFL-specific EFs, but in 2022, the same reservoir would have been considered FLRF and would have used FLRF EFs in that year.

For CO_2, emissions for each reservoir were estimated only for LCFL. Because CO_2 emissions were found to taper off after about a decade, the IPCC only provides guidelines for estimating CO_2 for recently flooded areas; reservoirs more than 20 years old were assumed to have no CO_2 emissions (Buendia *et al.*, 2019). Equation 1 was used to calculate CO_2 emissions per reservoir, denoted as $LCFL_{CO_2}$, using the EFs shown in Table 2, converted to tonnes m^{-1} y^{-1} from the IPCC (Buendia *et al.*, 2019).

Equation 1. Total CO_2 emissions for LCFL. A is the surface area of the reservoir, and $EF_{LCFL\,CO_2,j}$ is the emissions factor in Table 2, for the reservoir's climate zone, j, from Table 1.

$$LCFL_{CO_2} = A * EF_{LCFL\,CO_2,j}$$

Table 2. CO_2 emissions factors for LCFL based on the climate zone. The "n" column denotes the number of reservoirs used to calculate the standard deviation using Equation 5.

Climate Zones	j	LCFL Emissions Factors, tonnes CO_2 m^{-1} y^{-1}	Standard Deviation	n
Boreal	1	0.00035	0.00021	118
Cool temperate	2	0.00037	0.00017	2103
Warm temperate dry	3	0.00062	0.00022	679
Warm temperate moist	4	0.00054	0.00017	2095
Tropical dry/montane	5	0.0011	0.00051	902
Tropical moist/wet	6	0.0010	0.00037	920

CH_4 emissions estimates were calculated for both LCFL, in Equation 2, and FLRL, in Equation 3. The equations are identical except for differences in their EFs, shown in Table 3 and Table 4, respectively. Note that CH_4 emissions are influenced by the trophic state of the reservoir, which determines the value of α, the EF adjustment variable, which increases as a reservoir undergoes eutrophication. For our methodology, this value was always 1.0, based on the IPCC's default value for tier 1 emissions. CH_4 emissions in areas directly downstream of a reservoir were also considered in these equations. The variable R_d represents the ratio of downstream CH_4 emissions to the flux of CH_4 from the reservoir itself. Again, the default value of 0.09 was used for this ratio (Buendia *et al.*, 2019). Note that in our reported Climate TRACE data for the reservoirs sector, the CH_4 EFs column accounts for downstream emissions so that those EFs are 1.09 times higher than the values in Tables 2 and 3. Therefore, in the reported data, CH_4 emissions are simply the product of the EF and capacity columns.

Equation 2. Total CH_4 emissions for LCFL. A is the surface area of the reservoir, $EF_{LCFL\,CH_4,j}$ is the EF in Table 3 for the reservoir's climate zone, j, α is the emissions factor adjustment for trophic state, equal to 1.0, and R_d is the ratio of downstream CH_4 emissions to the flux of CH_4 from the reservoir, equal to 0.09.

$$LCFL_{CH_4} = LCFL_{CH_4,res} + LCFL_{CH_4,downstream}$$

$$LCFL_{CH_4,res} = \alpha\left(A * EF_{LCFL\,CH_4,j}\right)$$

$$LCFL_{CH_4,downstream} = \alpha\left(A * EF_{LCFL\,CH_4,j}\right) * R_d$$

Table 3. CH_4 emissions factors for LCFL based on the climate zone. The "n" column denotes the number of reservoirs used to calculate the standard deviation using Equation 5.

Climate Zones	j	LCFL Emissions Factors, tonnes CH_4 m^{-1} y^{-1}	Standard Deviation	n
Boreal	1	$2.77 * 10^{-6}$	$3.47 * 10^{-6}$	96
Cool temperate	2	$8.47 * 10^{-6}$	$1.30 * 10^{-5}$	1879
Warm temperate dry	3	$1.96 * 10^{-5}$	$2.32 * 10^{-5}$	578
Warm temperate moist	4	$1.28 * 10^{-5}$	$1.34 * 10^{-5}$	1946
Tropical dry/montane	5	$3.92 * 10^{-5}$	$3.48 * 10^{-5}$	710
Tropical moist/wet	6	$2.52 * 10^{-5}$	$2.18 * 10^{-5}$	805

Equation 3. Total CH_4 emissions for FLRF. A is the surface area of the reservoir, $EF_{FLRL\,CH_4,j}$ is the emissions factor in Table 4 for the reservoir's climate zone, j, α is the emissions factor adjustment for trophic state, equal to 1.0, and R_d is the ratio of downstream CH_4 emissions to the flux of CH_4 from the reservoir, equal to 0.09.

$$FLRF_{CH_4} = FLRF_{CH_4,res} + FLRF_{CH_4,downstream}$$

$$FLRF_{CH_4,res} = \alpha\left(A * EF_{FLRF\,CH_4,j}\right)$$

$$FLRF_{CH_4,downstream} = \alpha\left(A * EF_{FLRF\,CH_4,j}\right) * R_d$$

Table 4. CH_4 emissions factors for FLRF based on the climate zone. The "n" column denotes the number of reservoirs used to calculate the standard deviation using Equation 5.

Climate Zones	j	FLRF Emissions Factors, tonnes CH_4 m^{-1} y^{-1}	Standard Deviation	n
Boreal	1	$1.36 * 10^{-6}$	$3.15 * 10^{-6}$	96
Cool temperate	2	$5.40 * 10^{-6}$	$1.24 * 10^{-5}$	1879
Warm temperate dry	3	$1.51 * 10^{-5}$	$2.13 * 10^{-5}$	578
Warm temperate moist	4	$8.03 * 10^{-6}$	$1.35 * 10^{-5}$	1946
Tropical dry/montane	5	$2.84 * 10^{-5}$	$2.98 * 10^{-5}$	710
Tropical moist/wet	6	$1.41 * 10^{-5}$	$1.56 * 10^{-5}$	805

From CO_2 and CH_4 emissions, CO_2eq was calculated for each reservoir using Equation 4. The 20-year and 100-year global warming potentials (GWP) for CH_4 were taken from the IPCC's Sixth Assessment Report (2023).

Equation 4. CO2 equivalents for 20-year and 100-year GWPs. $GWP_{CH_4 20yr}$ is equal to 80.8 and $GWP_{CH_4 100yr}$ is equal to 27.2.

$$CO_2eq_{20yr} = CO_2 + GWP_{CH_4 20yr} * CH_4$$

$$CO_2eq_{100yr} = CO_2 + GWP_{CH_4 100yr} * CH_4$$

Uncertainty, reported as a standard deviation, was also estimated for each reservoir's surface area and for CO_2, CH_4, and for CO_2eq emissions and EFs. For each reservoir's surface area, the uncertainty was calculated as the standard deviation between the "Area_poly", "Area_rep", "Area_max", and "Area_min" columns in GRanD; if only one column had a reported value, the standard deviation was taken as the "Area_skm" value multiplied by the average percent standard deviation across the rest of the reservoirs. For CO_2 and CH_4 EFs, the standard deviations shown in Tables 2, 3, and 4 were calculated using Equation 5 to convert from the confidence intervals reported by the IPCC (Higgins, Li and Deeks, 2023). For CO_2 EFs, n was not directly reported by the IPCC and was assumed to be the same as those reported for tier 2 CO_2 scaling factors, since both seem to have been derived from the same model.

Equation 5. Standard deviation derived from confidence intervals. CI_{upper} and CI_{lower} are the upper and lower bounds of the confidence intervals for the value of interest. n is the number of values used to estimate the value of interest, assuming it is an average of several measurements. z is the z-score for a particular confidence interval, 1.96 for a 95% CI and 1.645 for a 90% CI.

$$\sigma = \frac{\left[CI_{upper} - CI_{lower}\right]}{2z} * \sqrt{n}$$

Table 5. Constants and their uncertainties. The "n" column denotes the value used to calculate the standard deviation using Equation 5.

Constant Name	Value	Standard Deviation	n
R_d	0.09	0.26	36
$GWP_{CH_4 20yr}$	80.8	85.9	30
$GWP_{CH_4 100yr}$	27.2	36.6	30

Uncertainty for CO_2, CH_4, and CO_2eq emissions values were estimated using 1 million Monte Carlo simulations. For CO_2eq, this required estimating the standard deviation for the GWP values using Equation 5, with n assumed to be 30. Because R_d was included in the reported CH_4 EFs, the Climate TRACE uncertainty data for CH_4 EFs also included the uncertainty of R_d. The same Monte Carlo method was used to combine these uncertainties using the standard deviations in Table 5.

Confidence for capacity, type, and all emissions and EFs were reported per-reservoir on a 5-point scale: very low, low, medium, high, very high. For capacity and type, these values were taken directly from GRanD's "Quality" column, which uses a similar 5-point scale. All EFs were assigned a medium confidence, Climate TRACE's default value for IPCC tier 1 EFs. For CO_2, CH_4, and CO_2eq emissions values, the confidence was medium, unless capacity had a lower confidence, in which case, that confidence value was used.

2.3 Verifying modeled emissions estimates

At the time of this writing, there were few other publications that used the IPCC's methodology for calculating tier 1 reservoir emissions. As such, it was difficult to find direct comparisons for validation. The IPCC emissions estimates seemed to be lower than previously reported values. However, there were a few case studies that used the same EFs with similar results; work from

Sánchez-Carrillo *et al.* (2022) and Chung *et al.* (2022), for example, contained estimates within the same magnitude as the ones presented here.

3. Results

Out of the 6,275 reservoirs in our dataset, about half are in temperate climate zones. A lower percentage of our reservoirs are assigned to tropical climate zones, compared with the IPCC's dataset. The global distribution of reservoirs is shown in Figure 2. For 2022 estimates, 42 reservoirs in our dataset are labeled as LCFL while the rest are FLRF, meaning that the vast majority of emissions are driven by CH_4 alone. Summing across all reservoirs, the total emissions for this sector in 2022 are 251,942,373 tonnes CO_2eq_{20yr}, composed of 3,089,489 tonnes of CH_4 and 2,311,643 tonnes of CO_2. Over time, reservoir emissions seem to be slowly decreasing. Emissions have dropped by 4 percent since 2015, with the largest annual decrease occurring between 2019 and 2020. Emissions dropped by nearly 2 percent between those two years alone, shown in Figure 3. This coincides with the steady decrease in LCFL-labeled reservoirs, shown in Figure 4. The largest annual decrease in LCFL capacity, around 38 percent, also occurred between 2019 and 2020.

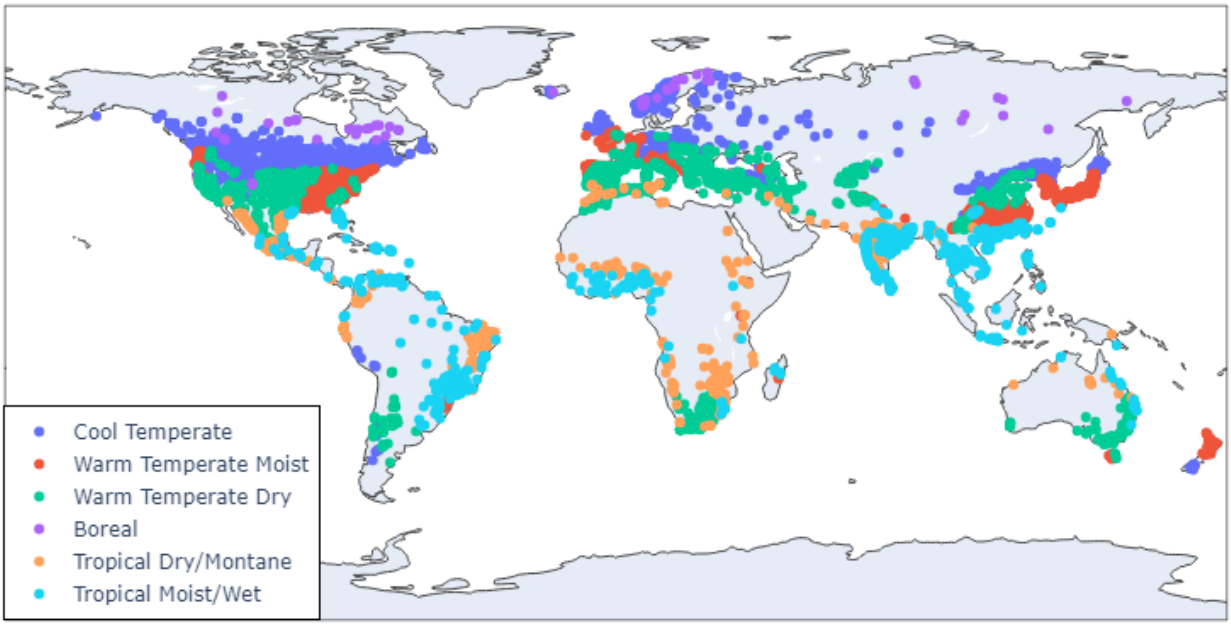

Figure 2. Locations of reservoirs in our dataset, with climate zones color-coded.

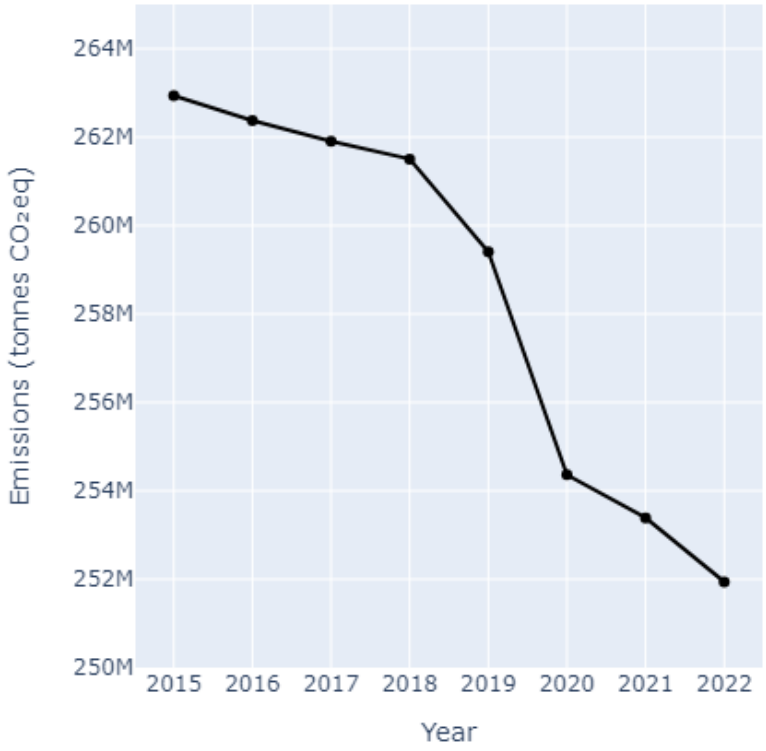

Figure 3. CO_2eq_{20yr} emissions from all reservoirs from 2015 to 2022.

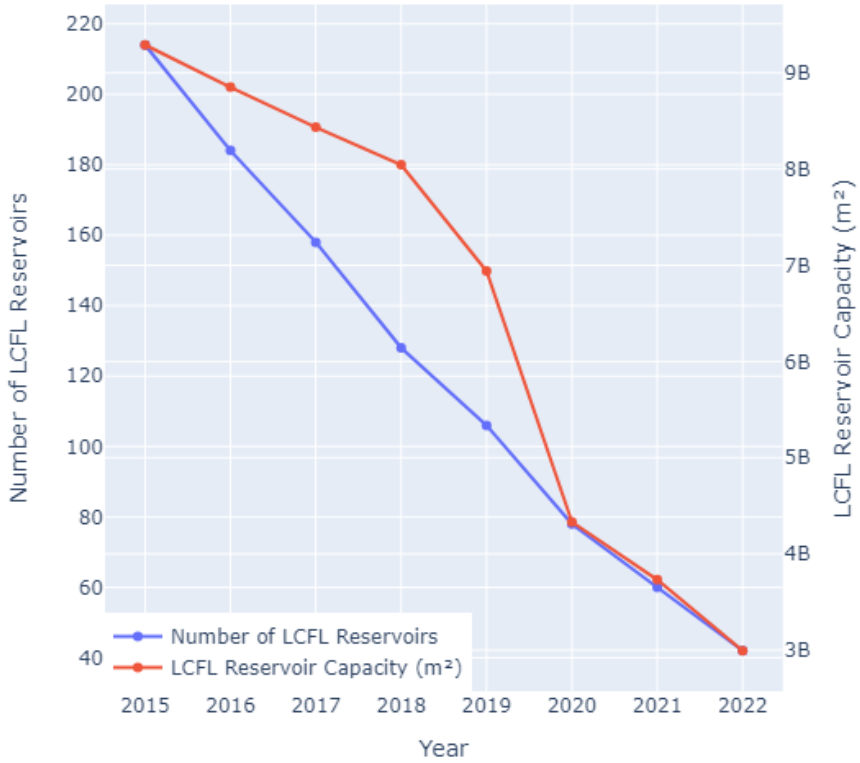

Figure 4. Number and capacity of LCFL-labeled reservoirs from 2015 to 2022.

Many of the reservoirs with the highest emissions, shown in Figure 5, are from tropical climate zones where EFs were largest. Lake Nasser, for instance, is located in Egypt in a tropical dry zone; it is created by Aswan High Dam and is the largest reservoir in the dataset by surface area. Lake Kariba, the largest artificial lake in the world by volume, located in Zimbabwe, is also in a tropical dry zone. Both of these reservoirs have no CO_2 emissions in 2022; however, their large capacity and CH_4 EFs mean that they have the highest total emissions. The countries with the highest emissions often also contain the largest combined reservoir capacities, although countries in cooler climate zones often have lower overall emissions, shown in Figure 6. For instance, Russia lies primarily in cool and boreal climate zones, which have lower EFs than warm and tropical zones. This means that, while Russia has the highest reservoir capacity of any country, its emissions are only the fourth highest overall. In contrast, the USA, which has a more even mix of reservoirs across warm and cold climate zones, has both the second highest capacity and the second highest resulting emissions.

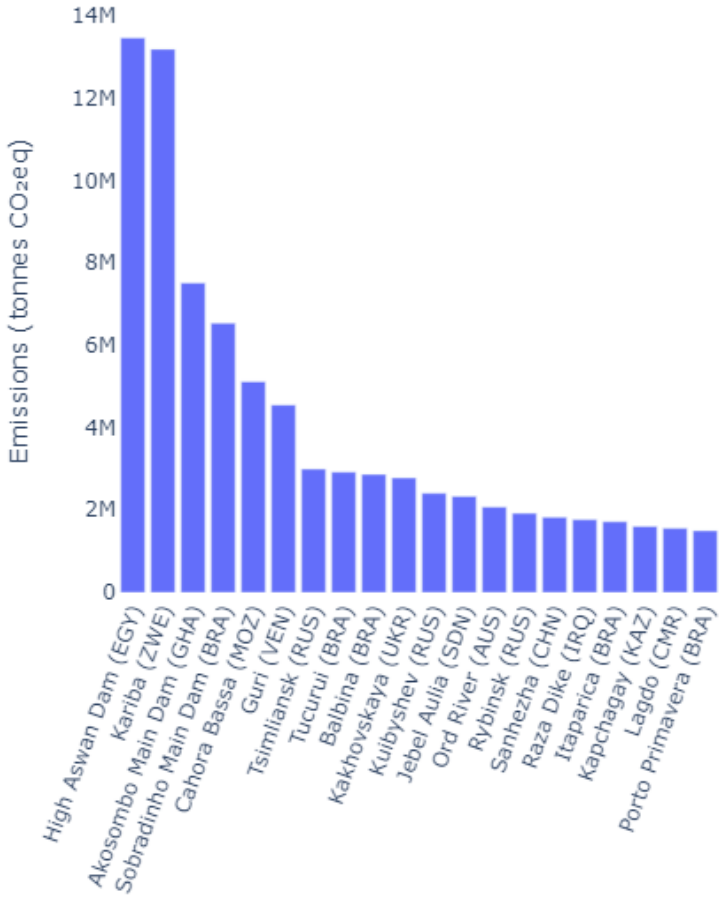

Figure 5. 2022 top 20 reservoir emissions in million tonnes CO_2eq_{20yr}.

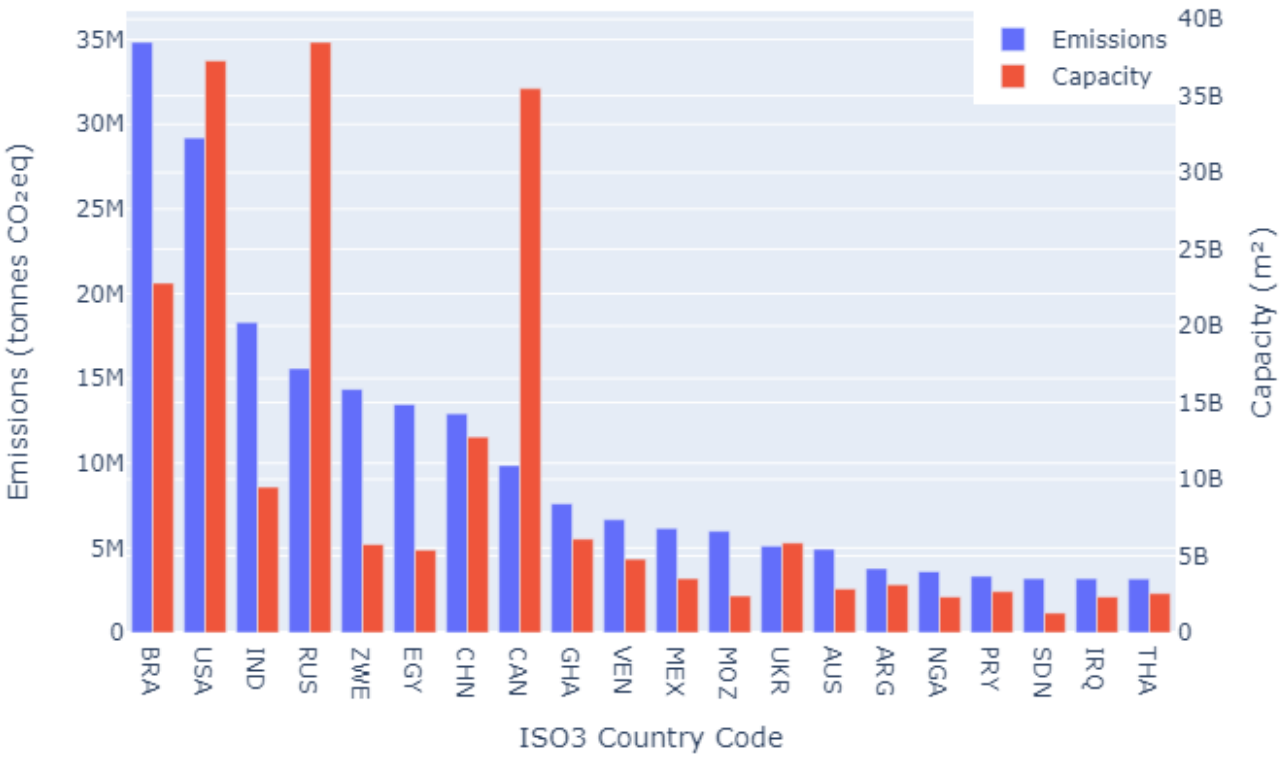

Figure 6. Comparison of total country emissions (left) and reservoir capacities (right).

Uncertainties for emissions values, shown in Figure 7, are high, driven primarily by large discrepancies in capacity. High uncertainty in capacity could partially be due to fluctuations in reservoir surface area as the water level is regulated via dams, causing large differences in the "Area_min" and "Area_max" columns in GRanD; discrepancies between "Area_rep" and "Area_poly" are also common in the dataset. Large standard deviations for constants and EFs also contribute to high overall uncertainty. CH_4 EFs have a much higher uncertainty than CO_2 EFs, partially due to the inclusion of the R_d value's uncertainty in its calculation. Once converted to standard deviations, both CO_2 and CH_4 EFs have uncertainties of similar magnitude to the values themselves. The IPCC states that EFs "represent global averages and have large uncertainties due to variability in climate and management practices" (Buendia *et al.*, 2019). For 2022, 7 percent of reservoirs have emissions values with low or very low confidences, again due to low capacity confidences, while the remaining 93 percent of reservoirs have medium confidences for emissions.

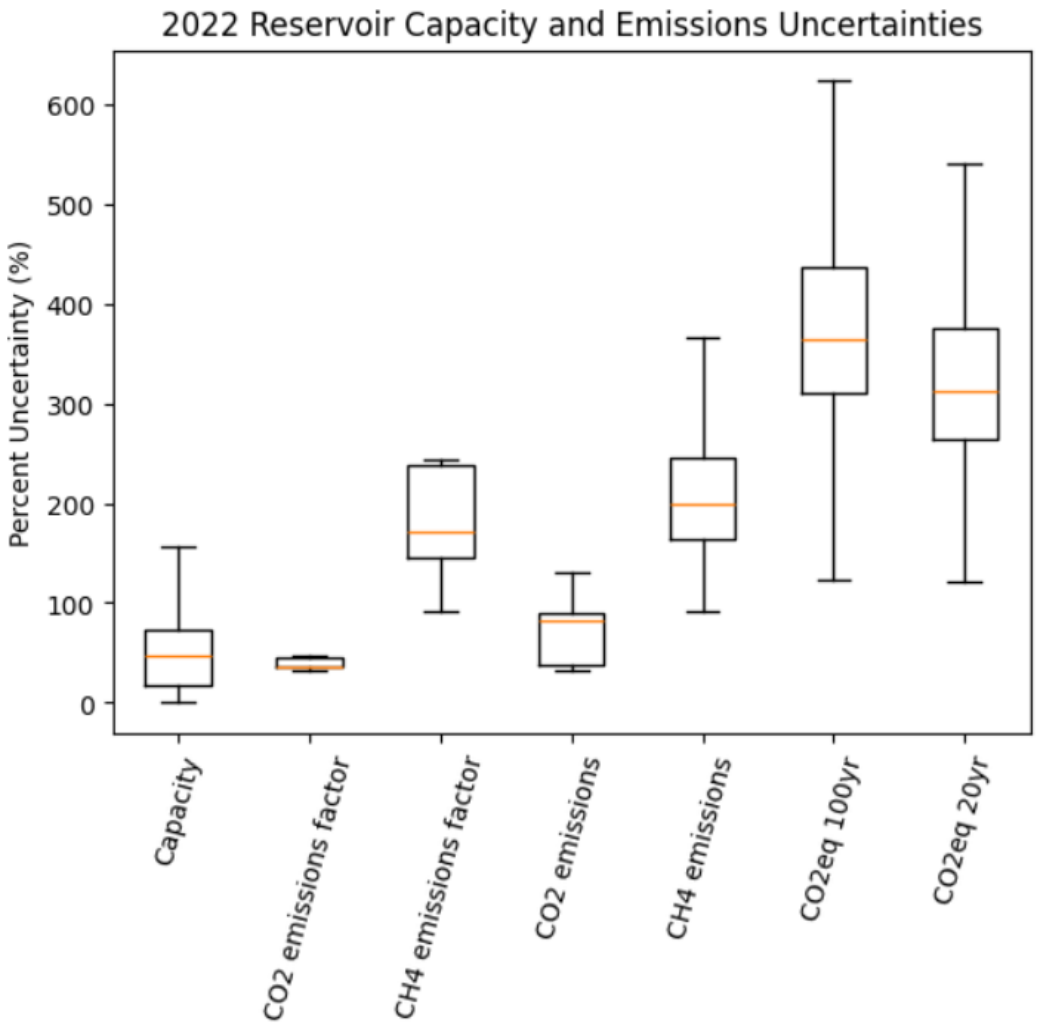

Figure 7. Reservoir capacity, EF, and emissions uncertainties for 2022, presented as percent standard deviations. Note that some reservoirs are outliers and are not on this graph, with deviations in the thousands of percent due to very high deviations in capacity.

4. Discussion

The emissions estimates calculated here are very conservative compared to other published estimates, shown in Table 6, despite using surface area and EF estimates that are similar in magnitude to other studies. There are a few improvements that could be made to increase accuracy and to include potential missing sources of anthropogenic emissions. Soued et al. (2022) were able to use the IPCC's EFs in the G-res model to create a total estimate of over 1 billion tonnes CO_2eq in 2020 with a similar set of reservoirs. Given this, it may be that the equations, rather than the EFs, for tier 1 emissions produce conservative numbers. For example, the vast majority of reservoirs have no CO_2 emissions in 2022 since they are labeled FLRF. It could be the case that these reservoirs do have lower, but non-zero emissions after 20 years, and it may be more accurate to gradually decrease emissions with age as a variable. For CH_4, the

default values used in Equations 2 and 3 could be adjusted per-reservoir, especially α. Eutrophic reservoirs, for instance, could have around 14 times the emissions of oligotrophic reservoirs due to a larger α value (Buendia *et al.*, 2019). It may also be more accurate to use the IPCC's original equations for CO_2 diffusive emissions and CH_4 diffusive and bubbling emissions, from which the EFs are derived (Buendia *et al.*, 2019). However, this would require collecting additional information such as solar irradiance and soil organic carbon for each reservoir.

Table 6. Emissions factors and annual estimates across previous studies, derived from Table 3 from Harrison et al. (2021).

Study	Surface Area Considered (m^2)	CO_2 EF, tonnes CO_2 m^{-1} y^{-1}	CH_4 EF, tonnes CH_4 m^{-1} y^{-1}	Estimated Emissions, tonnes CO_2eq_{20yr}
Harrison et al. (2021)	$3.50 * 10^{11}$	0.00094	$6.29 * 10^{-5}$	2,106,000,000
Deemer et al. (2016)	$3.11 * 10^{11}$	0.00044	$5.73 * 10^{-5}$	1,575,000,000
Hertwich (2013)	$3.30 * 10^{11}$	0.00085	$2.95 * 10^{-5}$	1,066,000,000
Basviken et al. (2011)	$3.40 * 10^{11}$	Not reported	$1.18 * 10^{-5}$	323,200,000
Barros et al. (2011)	$5.00 * 10^{11}$	0.00035	$4.00 * 10^{-5}$	1,792,000,000
St-Louis et al. (2000)	$1.50 * 10^{12}$	0.00067	$4.67 * 10^{-5}$	6,656,000,000

Another consideration is the dataset used. GRanD is often used for reservoir emissions estimates as its attribute data is complete and it includes the majority of large reservoirs. However, more extensive alternatives to GRanD, like the GOODD dataset (Mulligan, Soesbergen, and Sáenz, 2020), could be used to produce more extensive estimates if missing attribute data could be estimated. The IPCC also notes that, while many reservoirs that are formed from natural lakes would be excluded from anthropogenic emissions calculations, those reservoirs with significant changes in hydrology due to man-made activity should still be considered (Buendia *et al.*, 2019). We did not consider these reservoirs here, given that the anthropogenic component of emissions is difficult to separate from natural emissions, but they may be included in future, more extensive estimates.

5. Conclusion

The approach presented was an application of the IPCC's methodology to calculate anthropogenic GHG emissions and to estimate uncertainty for an extensive dataset of reservoirs. Water reservoirs emit both natural and anthropogenic GHG emissions; however, few open-source tools were available for separating these two emissions types. As such, only fully artificial reservoirs were included in our dataset, and their emissions estimates were low, although comparable in magnitude to those made by previous studies using the same EFs (Sánchez-Carrillo *et al.*, 2022; Chung *et al.*, 2022). Our results were also comparable to estimates created by Basviken et al., which only considered CH_4 emissions (2011). This highlighted an interesting side effect of the lower EFs used for FLRF reservoirs − because of the high number of FLRF reservoirs, many reservoirs had no CO_2 emissions in 2015 to 2022 using this methodology. CH_4 was the predominant gas contributing to emissions estimates, and as the rate of reservoir construction slows, it is likely that reservoir emissions will continue to decrease over time as existing reservoirs age into the FLRL category. The work of Soued et al. also showed decreasing reservoir emissions from 2017 onward, although the rate and magnitude of this decrease is an open research topic (2022).

The uncertainties in our emissions estimates were high, primarily due to discrepancies in officially reported and estimated surface area values in GRanD. While largely unaffected, these discrepancies also occasionally led to low confidence values for emissions estimates. In future estimates, this may be less of a concern as higher resolution reservoir boundaries, based on satellite imagery, become available in more recent reservoir datasets. While our estimates were not as extensive as those in previous studies, they can serve as a backbone for future improvements. The current methodology uses global averages for EFs, while ideally, future estimates would calculate EFs and net anthropogenic emissions by considering each reservoir's geographic and environmental attributes separately. As discussed, a reservoir's trophic state, climate zone, and soil organic carbon content, among other variables, can considerably change its greenhouse gas emissions. We plan to improve future estimates by using methodologies that take more reservoir attributes into consideration, such as the IPCC's tier 2 emissions equations, or a linear regression model that can generate equations based on correlations between several reservoir attributes and real-world emissions measurements. We also plan to create more complete capacity estimates by including smaller reservoirs and partially-natural reservoirs with their own EFs, separate from those used by fully human-made reservoirs.

6. Supplemental section metadata

Water reservoirs' emissions are reported for CH_4, CO_2, and 20- and 100-year CO2eq for 6,275 sources for years 2015-2022, along with associated confidences and uncertainties. Country-level emissions for 128 countries are also available, created by summing emissions for all reservoirs within each country. Note that only water reservoirs that are fully human-made are included in

the dataset; reservoirs created by modifying natural lakes with regulation structures, like dams, are not included in these estimates. More detailed information on emissions and EFs for included reservoirs is available in Tables S1, and S2.

Table S1 General dataset information for reservoirs.

General Description	Definition
Sector definition	*Water reservoirs*
UNFCCC sector equivalent	*4.D.1.b., 4.D.2.b.*
Temporal Coverage	*2015 – 2022*
Temporal Resolution	*Annual*
Data format(s)	*CSV*
Coordinate Reference System	*Coordinates of each reservoir given in decimal degrees*
Number of countries/sources available for download and percent of global emissions (as of 2022)	*128 countries total, representing 6,275 individual water reservoirs globally.*
Total emissions for 2022	*251,942,373 tonnes 20-year CO_2eq*
Ownership	*Country*
What emission factors were used?	*IPCC tier 1 EFs, Volume 4, CH.7, 2019rf*
What is the difference between a "NULL / none / nan" versus "0" data field?	*"0" values are for true non-existent emissions. If we know that the sector has emissions for that specific gas, but the gas was not modeled, this is represented by "NULL/none/nan"*
total_CO2e_100yrGWP and total_CO2e_20yrGWP conversions	Climate TRACE uses IPCC AR6 CO_2e GWPs. CO_2e conversion guidelines are here:https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_FullReport_small.pdf

Table S2 Source level metadata description confidence and uncertainty for reservoirs.

Data attribute	Confidence Definition	Uncertainty Definition
type	Converted from GRanD's quality score for the reservoir as follows: ● Very high: Verified (location and data have been verified) ● High: Good (location and data seem good but have not all been verified) ● Medium: Fair (some data discrepancies; missing data; or uncertainties) ● Low: Poor (significant data discrepancies of various kinds that indicate errors) ● Very low: Unreliable (severe data discrepancies without reasonable explanation)	N/A
capacity_description	Mirrors value for type	Standard deviation between following columns in GRanD: "Area_poly", "Area_rep", "Area_max", "Area_min". If only one of these columns is not null, product of "Area_skm" and the average percent standard deviation across all other reservoirs
capacity_units	m^2	m^2
capacity_factor_description	N/A	N/A
capacity_factor_units	N/A	N/A
activity_description	Mirrors value for capacity	Mirrors value for capacity
activity_units	N/A	m^2
CO2_emissions_factor	*Medium*: based on IPCC emissions factors	Taken from IPCC uncertainty estimates, converted to a standard deviation
CH4_emissions_factor	*Medium*: based on IPCC emissions factors	Combination of IPCC CH_4 EF and R_d uncertainty, expressed as a standard deviation
N2O_emissions_factor	N/A	N/A
other_gas_emissions_factor	N/A	N/A
CO2_emissions	*Medium*, unless source capacity is lower, in which case, same as capacity	Combination of IPCC CO_2 EF and capacity uncertainties,

		expressed as a standard deviation
CH4_emissions	*Medium*, unless source capacity is lower, in which case, same as capacity	Combination of IPCC CH_4 EF and capacity uncertainties, expressed as a standard deviation
N2O_emissions	N/A	N/A
other_gas_emissions	N/A	N/A
total_CO2e_100yrGWP	*Medium*, unless source capacity is lower, in which case, same as capacity	Combination of CO_2 EF, CH_4 EF, GWP and capacity uncertainties, expressed as a standard deviation
total_CO2e_20yrGWP	*Medium*, unless source capacity is lower, in which case, same as capacity	Combination of CO_2 EF, CH_4 EF, GWP and capacity uncertainties, expressed as a standard deviation

Data citation format: Brown, N., Volpato, G., Reilly, E. (2023). *Emissions from Water Reservoirs*. The Johns Hopkins University Applied Physics Laboratory (JHU/APL), USA, WattTime, USA, Climate TRACE Emissions Inventory. https://climatetrace.org [Accessed date]

Geographic boundaries and names (iso3_country data attribute): The depiction and use of boundaries, geographic names and related data shown on maps and included in lists, tables, documents, and databases on Climate TRACE are generated from the Global Administrative Areas (GADM) project (Version 4.1 released on 16 July 2022) along with their corresponding ISO3 codes, and with the following adaptations:

- HKG (China, Hong Kong Special Administrative Region) and MAC (China, Macao Special Administrative Region) are reported at GADM level 0 (country/national);
- Kosovo has been assigned the ISO3 code 'XKX';
- XCA (Caspian Sea) has been removed from GADM level 0 and the area assigned to countries based on the extent of their territorial waters;
- XAD (Akrotiri and Dhekelia), XCL (Clipperton Island), XPI (Paracel Islands) and XSP (Spratly Islands) are not included in the Climate TRACE dataset;
- ZNC name changed to 'Turkish Republic of Northern Cyprus' at GADM level 0;
- The borders between India, Pakistan and China have been assigned to these countries based on GADM codes Z01 to Z09.

The above usage is not warranted to be error free and does not imply the expression of any opinion whatsoever on the part of Climate TRACE Coalition and its partners concerning the

legal status of any country, area or territory or of its authorities, or concerning the delimitation of its borders.

Disclaimer: The emissions provided for this sector are our current best estimates of emissions, and we are committed to continually increasing the accuracy of the models on all levels. Please review our terms of use and the sector-specific methodology documentation before using the data. If you identify an error or would like to participate in our data validation process, please contact us.

References

1. Barros, N. *et al.* (2011) 'Carbon emission from hydroelectric reservoirs linked to reservoir age and latitude', *Nature Geoscience*, 4(9), pp. 593–596. Available at: https://doi.org/10.1038/ngeo1211.
2. Bastviken, D. *et al.* (2011) 'Freshwater Methane Emissions Offset the Continental Carbon Sink', *Science*, 331(6013), pp. 50–50. Available at: https://doi.org/10.1126/science.1196808.
3. Buendia, E. *et al.* (2019) *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*. The Intergovernmental Panel on Climate Change. Available at: https://www.ipcc-nggip.iges.or.jp/public/2019rf/pdf/4_Volume4/19R_V4_Ch07_Wetlands.pdf (Accessed: 31 October 2023).
4. Chung, Y., Paik, C. and Kim, Y.J. (2022) 'Estimation of Methane Emissions from Reservoirs Based on Country-Specific Trophic State Assessment in Korea', *Water*, 14(4), p. 562. Available at: https://doi.org/10.3390/w14040562.
5. Deemer, B.R. *et al.* (2016) 'Greenhouse Gas Emissions from Reservoir Water Surfaces: A New Global Synthesis', *BioScience*, 66(11), pp. 949–964. Available at: https://doi.org/10.1093/biosci/biw117.
6. Federici, S. (2021) *1st Corrigenda to the 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Available at: https://www.ipcc-nggip.iges.or.jp/public/2019rf/corrigenda1.html (Accessed: 31 October 2023).
7. Geffen, J. van (2023) *TEMIS -- GMTED2010 elevation data at different resolutions*. Available at: https://www.temis.nl/data/gmted2010/ (Accessed: 31 October 2023).
8. Harris, I. *et al.* (2020) 'Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset', *Scientific Data*, 7(1), p. 109. Available at: https://doi.org/10.1038/s41597-020-0453-3.
9. Harrison, J.A. *et al.* (2021) 'Year-2020 Global Distribution and Pathways of Reservoir Methane and Carbon Dioxide Emissions According to the Greenhouse Gas From Reservoirs (G-res) Model', *Global Biogeochemical Cycles*, 35(6), p. e2020GB006888. Available at: https://doi.org/10.1029/2020GB006888.
10. Hertwich, E.G. (2013) 'Addressing Biogenic Greenhouse Gas Emissions from

Hydropower in LCA', *Environmental Science & Technology*, 47(17), pp. 9604–9611. Available at: https://doi.org/10.1021/es401820p.
11. Higgins, J.PT., Li, T. and Deeks, J.J. (2023) *Cochrane Handbook for Systematic Reviews of Interventions version 6.4*, *Chapter 6: Choosing effect measures and computing estimates of effect*. Available at: https://training.cochrane.org/handbook/current/chapter-06 (Accessed: 31 October 2023).
12. Intergovernmental Panel on Climate Change (IPCC) (2023) *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/9781009157896.
13. Lehner, B. *et al.* (2011) 'High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management', *Frontiers in Ecology and the Environment*, 9(9), pp. 494–502. Available at: https://doi.org/10.1890/100125.
14. Messager, M.L. *et al.* (2016) 'Estimating the volume and age of water stored in global lakes using a geo-statistical approach', *Nature Communications*, 7(1), p. 13603. Available at: https://doi.org/10.1038/ncomms13603.
15. Mulligan, M., Soesbergen, A. van and Sáenz, L. (2020) 'GOODD, a global dataset of more than 38,000 georeferenced dams', *Scientific Data*, 7(1), p. 31. Available at: https://doi.org/10.1038/s41597-020-0362-5.
16. Prairie, Y.T. *et al.* (2021) 'A new modelling framework to assess biogenic GHG emissions from reservoirs: The G-res tool', *Environmental Modelling & Software*, 143, p. 105117. Available at: https://doi.org/10.1016/j.envsoft.2021.105117.
17. Sánchez-Carrillo, S. *et al.* (2022) 'Greenhouse gas emissions from Mexican inland waters: first estimation and uncertainty using an upscaling approach', *Inland Waters*, 12(2), pp. 294–310. Available at: https://doi.org/10.1080/20442041.2021.2009310.
18. Soued, C. *et al.* (2022) 'Reservoir CO2 and CH4 emissions and their climate impact over the period 1900–2060', *Nature Geoscience*, 15(9), pp. 700–705. Available at: https://doi.org/10.1038/s41561-022-01004-2.
19. St. Louis, V.L. *et al.* (2000) 'Reservoir Surfaces as Sources of Greenhouse Gases to the Atmosphere: A Global Estimate: Reservoirs are sources of greenhouse gases to the atmosphere, and their surface areas have increased to the point where they should be included in global inventories of anthropogenic emissions of greenhouse gases', *BioScience*, 50(9), pp. 766–775. Available at: https://doi.org/10.1641/0006-3568(2000)050[0766:RSASOG]2.0.CO;2.