Forestry and Land Use Change sector: Emissions from Reservoirs

Heather Hunter[1,3], Michael Robinette[1,3], Nicole Brown[1,3], Lekha Sridhar[2,3], Christy Lewis[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 historically 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). In 2023, 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. In 2024, regulated natural lakes with control structures were added to emissions calculations to generate more comprehensive emissions estimates.

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, emissions estimates across 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 every month between years 2015 to 2023 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), GRanD v1.3 (https://www.globaldamwatch.org/grand/; Lehner *et al.*, 2011), and GOODD (https://www.globaldamwatch.org/goodd; Mulligan *et al*., 2020), 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". In 2024, all reservoirs in GRanD that were part of the lake control category were included for anthropogenic emissions calculations. In addition, reservoirs from GOODD cross-listed as a "lake" in HydroLAKES were classified as regulated lakes and included for emissions estimates purposes. GOODD is a comprehensive global dataset consisting of over 38,000 dams (Mulligan *et al*., 2020). Reservoirs that had no estimated completion year, or that were listed as under construction, destroyed, or removed as of 2023 were also omitted from our dataset, leaving 7,184 reservoirs, 418 of which were classified as regulated natural lakes. Previously, 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. These data are static and represent surface areas at a single point in time. To support the expansion of emissions estimates to a monthly cadence in 2024, reservoir surface areas were derived from satellite imagery. Using up-to-date data enabled Climate TRACE to track surface area variability over time for better emissions estimates. To that end, surface areas were derived for all reservoirs in GRanD using a combination of Landsat and Sentinel-2 multispectral imagery acquired from Microsoft's Planetary Computer (https://planetarycomputer.microsoft.com/api/stac/v1) and Element84 (https://earth-search.aws.element84.com/), respectively, via the SpatioTemporal Asset Catalog (STAC) API. These two satellite missions represent the most widely and freely accessible multispectral satellite data at resolutions of 30-m and 10-m, respectively. When combined, data from Landsat and Sentinel-2 provide observations of every point on the globe every 2-3 days. Landsat is a series of joint NASA/USGS missions that has provided a continuous space-based record of Earth's surface since 1972. Sentinel-2, on the other hand, is a European Space Agency (ESA) mission that consists of two satellites which have been observing the Earth's surface continuously since 2015.

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 for December 2023. 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 based on the month of December 2023.

IPCC Climate Zones	Aggregated Climate Zones	j	n
Boreal dry	Boreal	1	1,828
Boreal moist			
Polar dry			
Polar moist			
Cool temperate dry	Cool temperate	2	2,634
Cool temperate moist			
Warm temperate dry	Warm temperate dry	3	446
Warm temperate moist	Warm temperate moist	4	283
Tropical dry	Tropical dry	5	1,069
Tropical montane			
Tropical moist	Tropical moist/wet	6	356
Tropical wet			

2.2 Methods

2.2.1 Deriving Climate Zones

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.

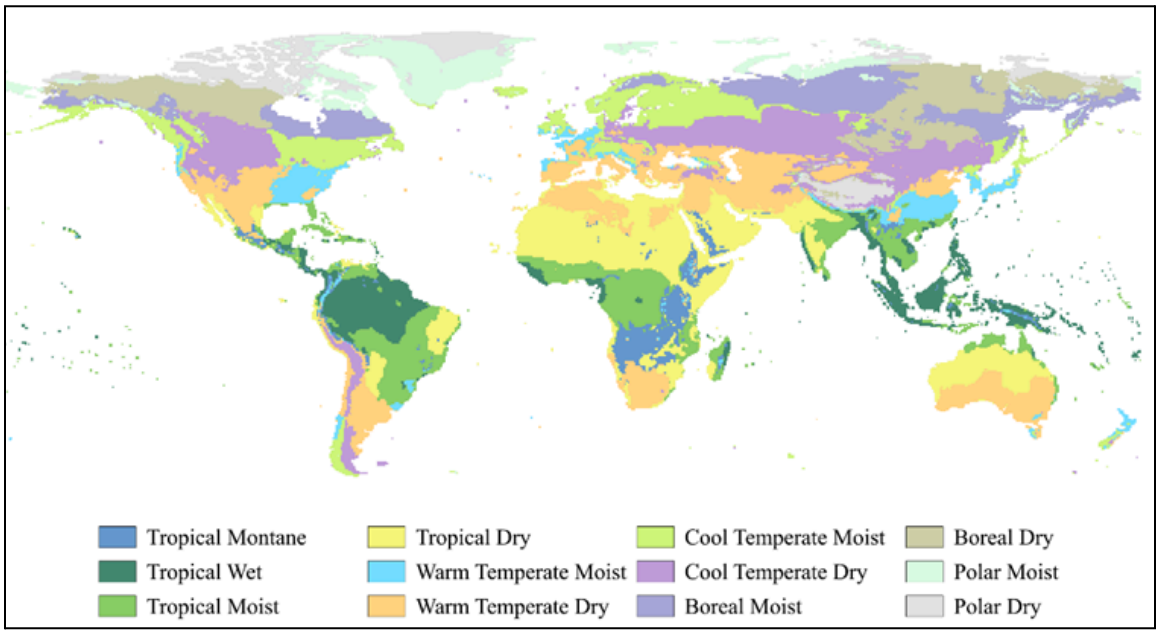

Figure 1. Recreated map of IPCC climate zones, with climate data from 1993-2022.

After being assigned a climate zone, reservoirs were labeled as FLRF or LCFL for each year from 2015 to July 2024. 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.

2.2.2 Deriving Reservoir Surface Area

To estimate the surface area of a given reservoir for a particular month, Sentinel-2 and Landsat data were queried from Element84 and Microsoft's Planetary Computer, respectively, using the STAC API, over the preceding 3-month period given a region of interest defined by the geographical boundaries in GranD. Only images with less than 30% cloud cover were queried. For Sentinel-2, bands 3 and 8 (green and near infrared) were queried, and for Landsat, different bands were queried depending on sensor availability at a given point in time. For example, Landsat 7's green band was band 2 and its infrared band was band 4 (see Equation 1), while Landsat 8's green band is band 3 and its infrared band is band 5 (see Equation 2). Landsat 7 operated from April 1999 to January 2024 and is the dominant source of imagery in this dataset.

For months following January 2024, Landsat 8 was used. For any region of interest that lacked Landsat data between 2015 and July 2024, Sentinel-2 data were used (see Equation 3). 3-month median composite images were then calculated for the green and near-infrared bands of each sensor. This means, for example, that a single near-infrared image was generated for March 2023 where each pixel (x_i, y_i) was set to the median value of all corresponding pixels (x_i, y_i) over every near-infrared image from January 2023, February 2023, through March 2023. This is common practice in the field of remote sensing to mitigate the effects of cloud cover and other short-lived, non-persistent activity (e.g., anthropogenic activities, severe weather), particularly in areas of interest that are slowly varying, like reservoirs.

To monitor spatiotemporal changes in water bodies imaged by multispectral sensors, it is common practice to use the fast and simple approach of the Normalized Difference Water Index (NDWI) (McFeeters, 1996). The NDWI is a spectral index calculated using multispectral imagery in the green and near-infrared portions of the electromagnetic spectrum, which means it can be applied to any sensor that acquires data in these bands. It is generally good at enhancing water bodies due to the relatively strong reflectance of water in the green band and water's absorption of light in the near-infrared, though this can depend on the clarity (or turbidity) of the water and the depth of the water body.

Equations 1, 2, or 3 were used to calculate the NDWI for each monthly median composite image, depending on imagery source. As denoted in Table 1, a basic decision tree was then used to classify the type of pixel associated with a particular value of NDWI. Surface areas were then computed as the summation of all NDWI pixels greater than or equal to 0.2. If the estimated surface area was zero, the reservoir was assumed to be ice- or snow-covered if the climate zone for that reservoir was boreal. If the climate zone for that reservoir was not boreal, the reservoir pixels were likely corrupted either by clouds or sensor processing issues (e.g., dropped pixels or striping). For boreal regions, it was assumed that ice covered reservoirs had zero emissions; thus, a surface area of zero was valid, as that ensured the reservoir would not contribute to overall emissions estimates.

Equation 1. NDWI (McFeeters, 1996) for Landsat 7. *R* refers to measured reflectance values at the satellite sensor within a given band, e.g., green (band 2) and near-infrared (band 4).

$$NDWI = \frac{R_{B02} - R_{B04}}{R_{B02} + R_{B04}}$$

Equation 2. NDWI for Landsat 8. *R* refers to measured reflectance values at the satellite sensor within a given band, e.g., green (band 3) and near-infrared (band 5).

$$NDWI = \frac{R_{03} - R_{05}}{R_{03} + R_{05}}$$

Equation 3. NDWI for Sentinel-2. *R* refers to measured reflectance values at the satellite sensor within a given band, e.g., green (band 3) and near-infrared (band 8).

$$NDWI = \frac{R_{03} - R_{08}}{R_{03} + R_{08}}$$

Table 1. NDWI ranges and their corresponding pixel types. Different ranges of NDWI correspond to different types of surfaces and indicate the presence or absence of water in some form (i.e., liquid, ice).

NDWI Value	Pixel Type
0.2 – 1.0	Water surface
0.0 – 0.2	Ice or snow cover (if climate zone is boreal), clouds, vegetation
-1.0 – 0.0	Dry land, non-aqueous surfaces

If surface area estimates were zero or indeterminate for a given month, but non-zero for other months, the surface area estimate from the month nearest in time was used. In this scenario, there were several months where a reservoir's surface area did not appear to change. If, due to a lack of sufficient satellite data coverage, surface area estimates were indeterminate for an entire year, monthly surface estimates were simply set to the default surface area value in GRanD.

Note that while the NDWI provides a sufficient first estimate for water body extraction, it can underestimate water body pixels due to its sensitivity to persistent cloud cover such as often seen over boreal regions (i.e., Canada) and the turbidity of water. This means that surface areas for water bodies that are cloud covered for most of the year, shallow, or inundated with run-off associated with human activities or natural factors (precipitation, severe weather) may be underestimated.

2.2.2 Estimating Reservoir Emissions

Estimating the anthropogenic emissions of regulated natural lakes required the estimation of the proportion of the lake that was man-made. To accomplish this, the difference between the surface area prior to the construction of the control structure was subtracted from present day surface area. Because the surface area prior to construction was unknown for most of these lakes, historical satellite imagery was collected to provide these estimates. While Landsat is the best source of global historical satellite imagery, its global record at 30-m spatial resolution only

dates back to 1982. Thus, Climate TRACE selected lakes whose control structures were built after 1982. Landsat imagery was queried over a 6-month period prior to the construction date of the control structure. Given the resulting 6-month median image, the surface area of the lake (or river, in some cases) was calculated by summing water pixels identified by the NDWI. This surface area was then subtracted from the corresponding monthly surface areas estimated from 2015 to July 2024. In total, 418 regulated natural lakes were identified in the GRanD database. 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 4 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). As shown in Equation 4, these EFs were multiplied by a decay parameter, δ, which was calculated using Equation 5 and derived from Prairie *et al.*, 2021. This decay parameter estimates the proportion of CO_2 emissions that changed since the reservoir was first created (initialized at $t = 0.5$ years). It is a time-varying value that considers the steady decrease in CO_2 emissions from LCFL from 0 years of age to 20 years of age. Its calculation is dependent on the mean annual air temperatures for each climate zone as well as the soil organic carbon content of a given reservoir, in addition to the reservoir's age, surface area, and total phosphorus (TP) content. At this time, due to the lack of time-varying measurements of TP and soil organic carbon content in reservoirs across the globe, these values were set to static values; as a result, because Equation 4 represents a division of exponents, these values are removed, and only the time-varying components (i.e., age, surface area) contribute to δ.

Equation 4. 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.

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

Equation 5. CO2 emissions decay parameter. *Y* is the age of the reservoir in years, $T_{eff}^{CO_2}$ is the effective air temperature in degrees Celsius, *A* is the surface area of the reservoir, *C* is the reservoir surface soil carbon content, and *TP* is the total phosphorus content.

$$\delta(t) = \frac{EM(t=0.5)}{EM(t)}$$

$$EM(t) = 10^{y}$$

$$y = 1.860 - 0.330Y + 0.332T_{eff}^{CO_2} + 0.0799A + 0.0155\,C + 0.2263(TP)$$

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	2,103
Warm temperate dry	3	0.00062	0.00022	679
Warm temperate moist	4	0.00054	0.00017	2,095
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 6, and FLRL, in Equation 7. 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 6. 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	$8.47 * 10^{-6}$	$1.30 * 10^{-5}$	1,879
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 7. 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 8. The 20-year and 100-year global warming potentials (GWP) for CH_4 were taken from the IPCC's Sixth Assessment Report (2023).

Equation 8. CO_2 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 reservoirs using the default GRanD 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 reservoirs with satellite-derived surface areas, uncertainty was set to 5%, which accounts for the average 3% radiometric uncertainty of reflectance measurements of the sensors from which the surface area is derived. For CO_2 and CH_4 EFs, the standard deviations shown in Tables 2, 3, and 4 were calculated using Equation 7 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 9. 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 9, 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 capacity factor, if a reservoir's surface areas were the default GRanD surface areas, the confidence was set to "low". If a reservoir's surface area was satellite-derived, the capacity and capacity factor confidence were set to "medium". 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 7,184 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 2023 estimates, 396 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 2023 are 307,948,517 tonnes CO_2eq_{20yr}, composed of 3,773,288 tonnes of CH_4 and 3,066,791 tonnes of CO_2. Through July 2024, the total emissions are

143,405,313 tonnes CO_2eq_{20yr}, composed of 1,754,958 tonnes of CH_4 and 1,604,669 tonnes of CO_2. Over time, reservoir emissions seem to be trending downward, with some variability. This coincides with the steady decrease in LCFL-labeled reservoirs, shown in Figure 4. Note in this figure the distinctly low capacities from 2015 to 2017. At this time, monthly satellite-derived surface area estimates are missing for these years, and surface area trends across these years are estimated using the static surface area values in GRanD. In the next iteration of this dataset, these values will be replaced with satellite-derived surface area estimates.

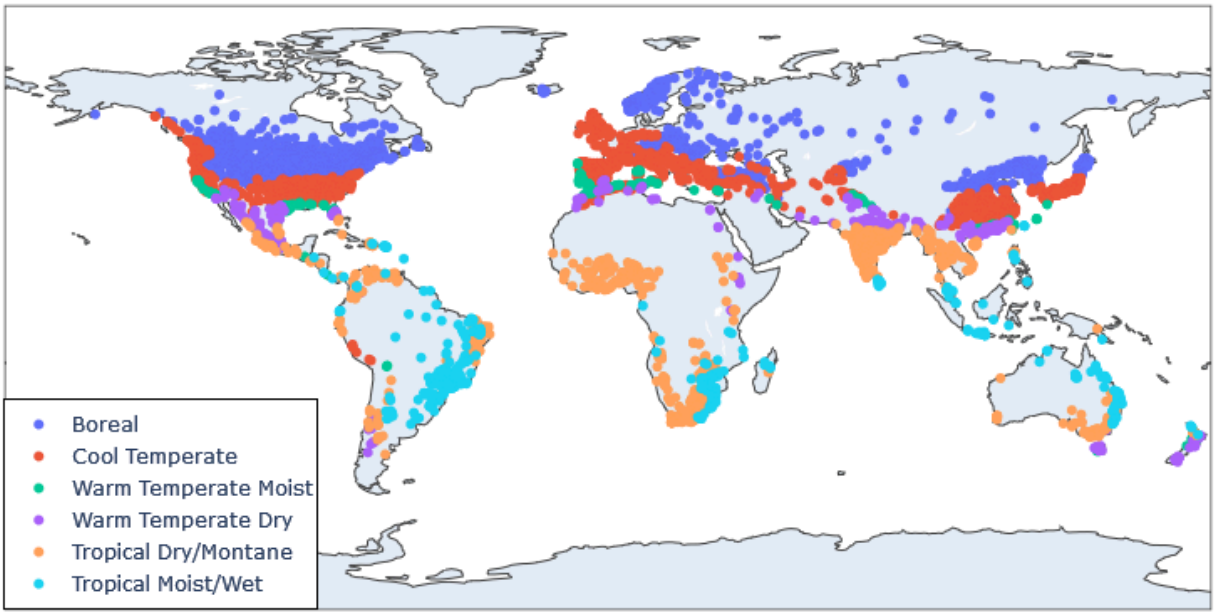

Figure 2. Locations of reservoirs in our dataset, with climate zones color-coded.

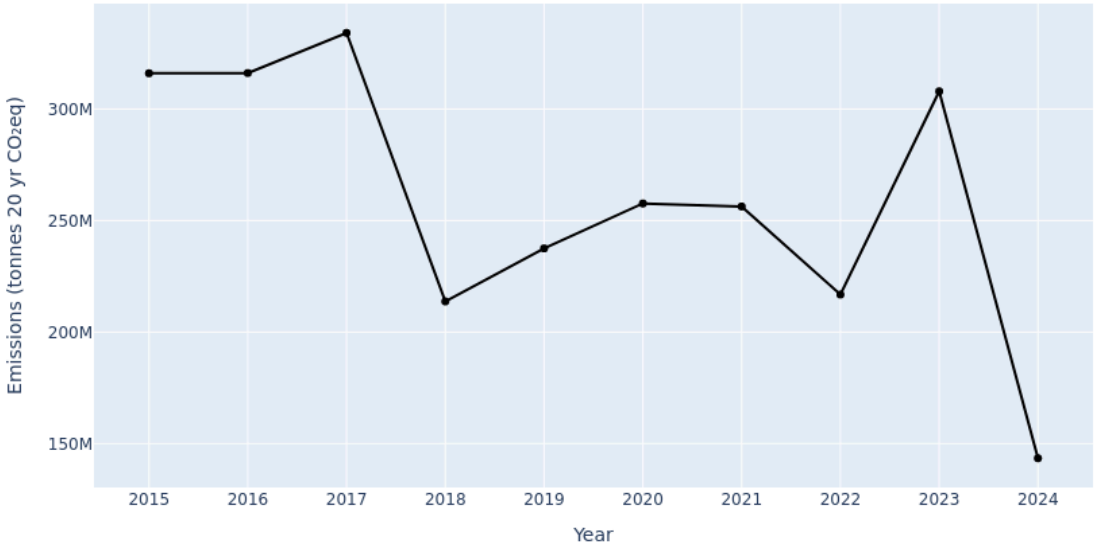

Figure 3. Total global CO_2eq_{20yr} emissions from all reservoirs from 2015 through July 2024.

Figure 4. Number and capacity of LCFL-labeled reservoirs from 2015 through July 2024.

Many of the reservoirs with the highest emissions in 2023, 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 High Aswan 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 2023; 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, Canada 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 fifth highest overall. In contrast, the USA, which has a more even mix of reservoirs across warm and cold climate zones, has slightly less capacity but the third highest resulting emissions.

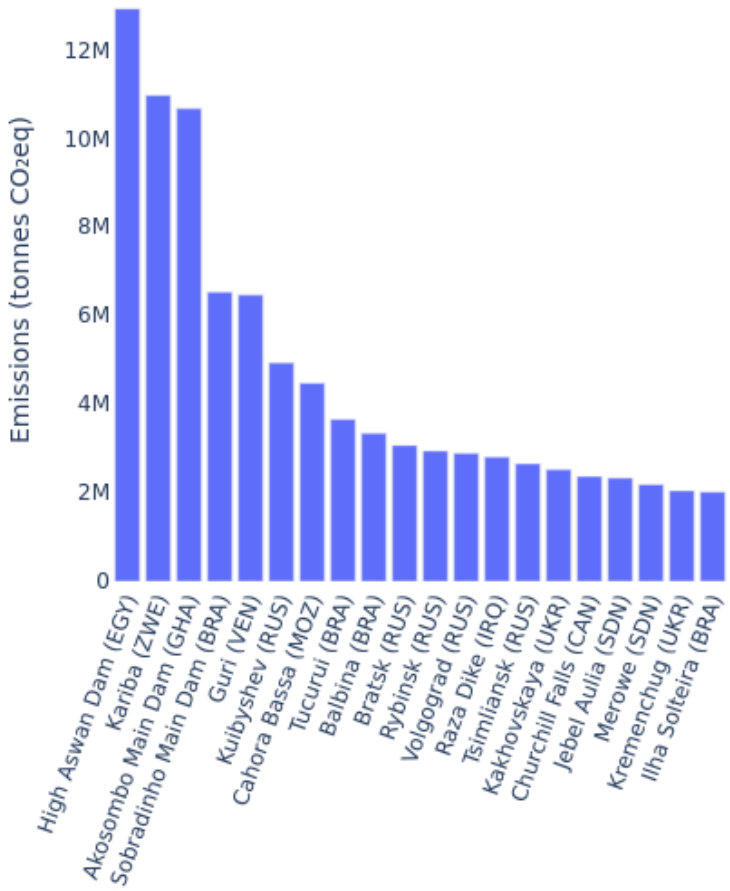

Figure 5. 2023 top 20 reservoir emissions in million tonnes CO_2eq_{20yr}.

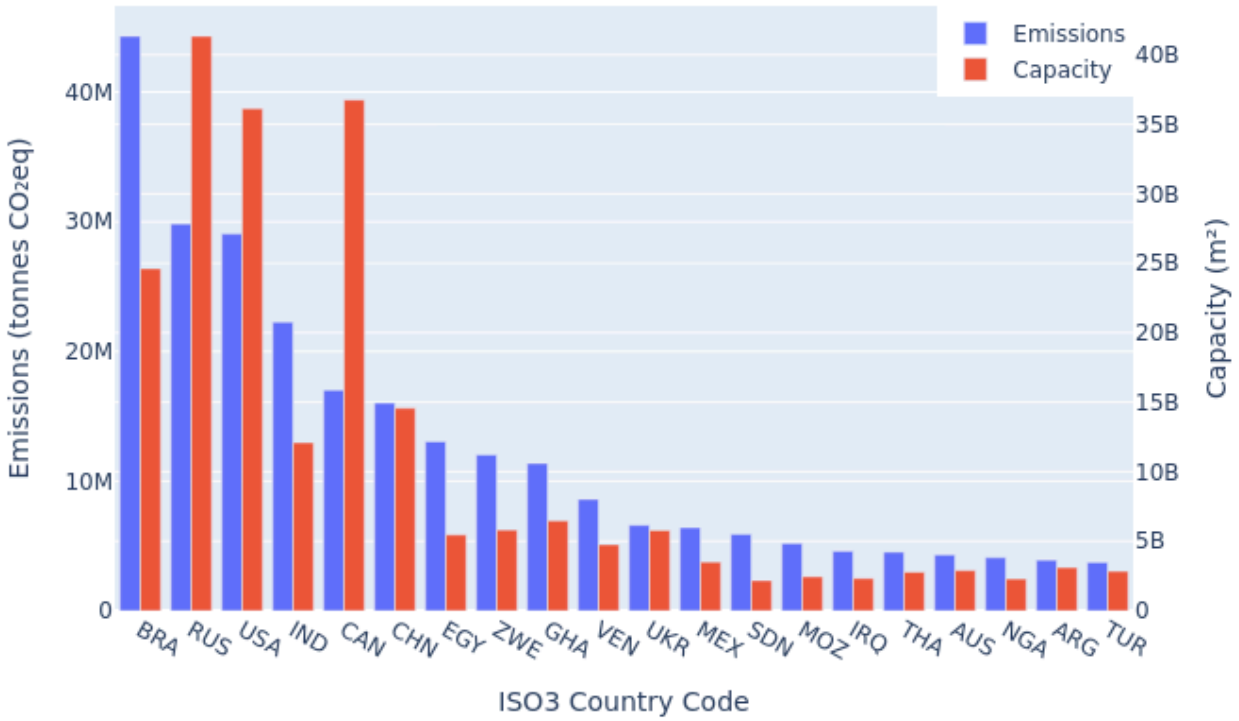

Figure 6. Comparison of total country emissions (left) and reservoir capacities (right) for 2023.

Uncertainties for emissions values, shown in Figure 7, are generally high, driven primarily by large discrepancies in CH_4 emissions and EFs. 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. In addition, the addition of the decay parameter described by Equation 3 drives the CO_2 EF lower overall. 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).

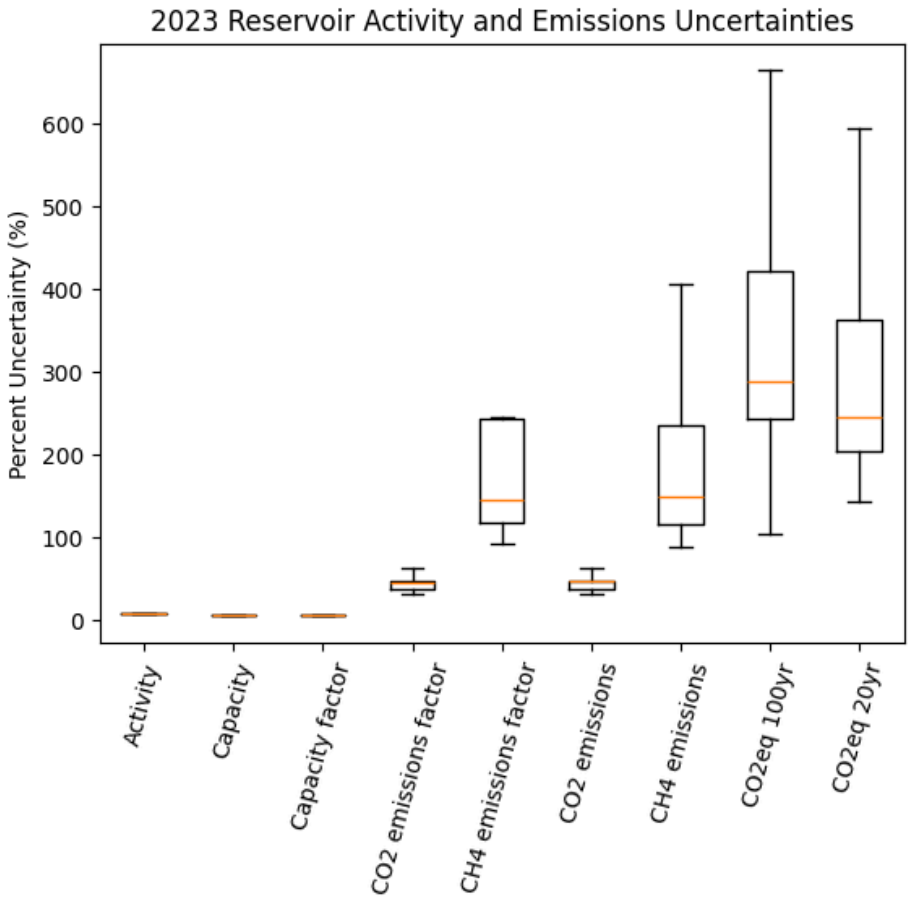

Figure 7. 2023 reservoir activity and emissions uncertainties, presented as percent standard deviations.

4. Discussion

The emissions estimates calculated here are very conservative compared to other published estimates, shown in Table 6, despite using updated, satellite-derived surface areas as well as 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 2023 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 beyond 20 years. While we do implement a CO_2 emissions decay equation for years 0 through 20, we will consider implementing it for later years as well. 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). While we began to consider these reservoirs here, in future iterations of this dataset we hope to expand the number of these regulated natural lakes.

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. Critically, we include both fully artificial reservoirs and regulated natural lakes in our dataset. Emissions from the artificial reservoirs 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 between 2015 and July 2024 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, driven primarily by uncertainties in CH_4 emissions factors. 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 through July 2024, along with associated confidences and uncertainties. Country-level emissions for 251 countries are also available, created by summing emissions for all reservoirs within each country. Note that water reservoirs that are both fully human-made and those that are created by modifying natural lakes with regulation structures, like dams, are included in these estimates. More detailed information on emissions and EFs for included reservoirs is available in Tables S1, S2, and S3. All of the data in these tables is available on the Climate TRACE website.

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	*January 2015 – July 2024*
Temporal Resolution	*Monthly*

General Description	Definition
Data format(s)	*CSV*
Coordinate Reference System	*Coordinates of each reservoir given in degrees*
Number of countries/sources available for download and percent of global emissions (as of July 2024)	*251 countries total, representing 7,184 individual water reservoirs globally.*
Total emissions for 2023	*307,948,517 tonnes 20-year CO_2eq*
Total emissions for January 2024 through July 2024	*143,405,313 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 for reservoirs.

Data attribute	Definition
sector	**water-reservoirs**
source_sub-sector_name	**N/A**
source definition	Individual water reservoir
start_date	Inclusive date of the emissions inventory. (YYYY-MM-DD)
end_date	Inclusive date of the emissions inventory. (YYYY-MM-DD)
source_identifier	Internal identifier of the water reservoir. Same as GRanD ID
source_name	Dam name according to GRanD. Reservoir name used where dam name was not available. If neither available, listed as "Water reservoir"
iso3_country	Iso3 country code, according to GRanD
location	WKT_point_location of the water reservoir, derived from centroid coordinates listed in GRanD
type	Reservoir main purpose, according to GRanD. If not available, listed as "other"

Data attribute	Definition
capacity_description	Reservoir surface area
capacity_units	m^2
capacity_factor_description	Constant with value 1 to match other sector definitions
capacity_factor_units	N/A
activity_description	Copy of capacity column to match other sector definitions
activity_units	m^2
CO2_emissions_factor	IPCC tier 1 carbon EF, specific to reservoir climate zone and age, tonnes CO_2 m^{-1} y^{-1}
CH4_emissions_factor	1.09 times IPCC tier 1 methane EF, specific to reservoir climate zone and age, tonnes CH_4 m^{-1} y^{-1}
N2O_emissions_factor	N/A
other_gas_emissions_factor	N/A
CO2_emissions	Estimated carbon emissions in tonnes CO_2
CH4_emissions	Estimated methane emissions in tonnes CH_4
N2O_emissions	N/A
other_gas_emissions	N/A
total_CO2e_100yrGWP	Total CO2e in tonnes using the CO_2e 100yr GWP conversion guidelines
total_CO2e_20yrGWP	Total CO2e in tonnes using the CO_2e 20yr GWP conversion guidelines

Table S3 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

Data attribute	Confidence Definition	Uncertainty Definition
capacity_description	Mirrors value for type. If the capacity is satellite-derived, confidence is *medium*. Otherwise, *low*.	If no satellite-derived surface area estimate exists: 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. For satellite-derived surface areas, uncertainty is set to ±5%.
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

Data attribute	Confidence Definition	Uncertainty Definition
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

Permissions and Use: All Climate TRACE data is freely available under the Creative Commons Attribution 4.0 International Public License, unless otherwise noted below.

Data citation format: Hunter, H., Robinette, M., Brown, N., Sridhar, L., Lewis, C., Reilly, E. (2024). *Forestry and Land Use Change sector- Emissions from 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.
20. McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425–1432. https://doi.org/10.1080/01431169608948714