

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Data of urban river greenhouse gas concentrations and fluxes and associated water physiochemical properties were either extracted directly from text and tables or estimated following the Fick's Law using local water physiochemical parameters as drivers. GetData Graph Digitizer (version 2.22) was used to extract data from figures. Reach-level catchment environmental and socioeconomic predictors were calculated by overlapping the catchments with relevant geospatial datasets using the mask tool of the Python Rasterio toolsets (version 1.4).
Data analysis	Data analysis was conducted with the software Microsoft Excel (version 2021), OriginPro (version 2023), randomForest package in R (version 4.2.1), and ArcGIS (version 10.8). The code used in this study is available from https://doi.org/10.6084/m9.figshare.24233902 .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The global morphological urban areas product is available from <https://data.mendeley.com/datasets/v3p8gk5724/1>. The Global Reach-scale A priori Discharge Estimates for SWOT (GRADES) dataset is available from <https://www.reachhydro.org/home/records/grades>. The Global River Methane Database (GriMeDB) is available from <https://doi.org/10.6073/pasta/f48cdb77282598052349e969920356ef>. The temperature and precipitation are available from WorldClim (version 2) (<https://www.worldclim.org/data/worldclim21.html>). The elevation and slope are available from Global Multi-resolution Terrain Elevation Dataset (<https://www.earthenv.org/topography>). The GDP and GDP per capita are available from Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015 (<https://datadryad.org/stash/dataset/doi:10.5061/dryad.dk1j0>). The population density is available from Gridded Population of the World (version 4) (<https://beta.sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density>). The MODIS gross and net primary productivity data are available from <https://www.umd.edu/numerical-terradynamic-simulation-group/project/modis/mod17.php>. The soil respiration rates data is available from <http://cse.affrc.go.jp/shojih/data/index.html>. Detailed information on an array of spatially explicit geospatial datasets used in this analysis is summarized in Supplementary Table 4. The dataset of urban river GHG concentrations and fluxes and related physiochemical properties is available from <https://doi.org/10.6084/m9.figshare.24233902>. Source data are provided with this paper.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

N/A

Population characteristics

N/A

Recruitment

N/A

Ethics oversight

N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

In this analysis, we aim to identify unique greenhouse gas emission profiles from global urban rivers, the environmental and socioeconomic drivers for their patterns and implications of the emissions. In order to achieve these goals, we compiled a data involving as many measurements of greenhouse gas concentrations/fluxes in urban rivers from the published literature. We then first reviewed present urban river greenhouse gas measurements in the literature and pointed out distinctive greenhouse gas emission profiles (e.g., particularly higher CH₄ and N₂O emissions) from urban than in non-urban rivers. Then, relying on a machine learning-based algorithm, we established predictive models between urban river greenhouse gas emissions and a range of reach-level catchment environmental and socioeconomic drivers, which allowed us to extrapolate our results to the global scale. These results finally allowed us a quantitative assessment of greenhouse emissions from global urban rivers, their geographic patterns, socioeconomic drivers and implications.

Research sample

The dataset of urban river greenhouse gas concentrations and fluxes was compiled from 116 published literature and contained 829, 809, and 644 for CO₂, CH₄, and N₂O concentrations, respectively; and 1,034, 914, 860 for CO₂, CH₄, and N₂O fluxes, respectively. Prediction of greenhouse gas emissions from urban rivers cover 1,554 global morphological urban areas.

Sampling strategy

Sampling strategy is not relevant to our study. All data were from published literature and open datasets.

Data collection

All data are from published literature and open datasets.

Timing and spatial scale

Data were compiled from the literature published between 1999 and 2023. Data includes greenhouse gas measurements in urban

Timing and spatial scale	rivers in Asia, Africa, Europe, North and South America and Oceania.
Data exclusions	We retained only studies that report in situ measurements of urban river GHG concentrations or fluxes. For studies that have measurements from both urban and non-urban river segments, all sites were retained for comparative analysis. We also kept the database strictly spatially explicit so that exact geographic locations of the measurement sites can be determined.
Reproducibility	Results can be reproduced following the data and methods described in the manuscript.
Randomization	Data were randomly split for training and testing of the random forest models.
Blinding	Blinding is not relevant to this study. All data were from published literature and open datasets.
Did the study involve field work?	☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging