Text S1 The data sources and modeling methods for machine learning.

The generation of waste were obtained from environmental statistics published form United Nations Statistics Division (https://unstats.un.org/unsd/envstats/country_files), and "What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050" published by World bank group (https://datacatalog.worldbank.org/search/dataset/0039597). To predict the global generation of waste, we prepared 10 gridded covariates. The detail information of regional income level, population, population density, gross domestic product (GDP), regional area, human development index (HDI), international annual tourists social progress index, social progress index, gross national income (GNI), and urban function were obtained from OECD data explorer (https://data-explorer.oecd.org/). The information of sewage discharge was obtained from United Nations Water (https://sdg6data.org/zh-hans/indicator/6.3.1). To predict the global generation of sewage, we prepared 8 gridded covariates. The information of population, population density, GNI per capita, HDI, GDP, HDI, and regional area were obtained from OECD data explorer (https://dataexplorer.oecd.org/). The average precipitation in depth and renewable internal freshwater resources per capita were collected from WorldClim (https://www.worldclim.org/).

A random forest regression combined with 10-fold cross validation was used for machine learning. To enhance the model's generalization ability, 10-fold cross-validation was used. We regenerated the data for machine learning relying on the factor loading results. The reliability of the waste generation model and the calculated feature importance are shown in Fig. S1 and Table S1. The reliability of the sewage discharge model and the calculated feature importance are shown in Fig. S2 and Table S2. Afterwards, the 49113 remaining data points were available for prediction approach, followed by visualizing the waste generation and sewage discharge worldwide via R package "ggplot2", "maps", and "dplyr".