Science of the Total Environment 872 (2023) 161913

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Socio-economic development drives solid waste management performance in cities: A global analysis using machine learning

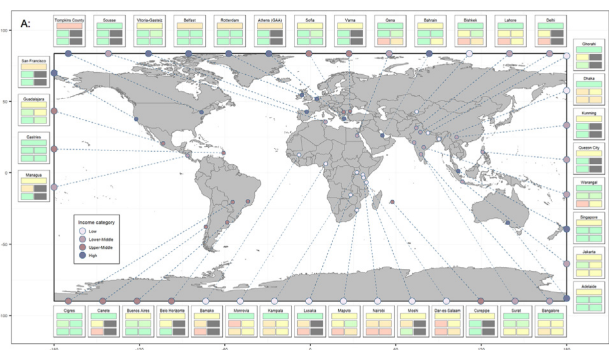
Costas A. Velis [a,*], David C. Wilson [b], Yoni Gavish [a], Sue M. Grimes [b], Andrew Whiteman [c]

[a] School of Civil Engineering, University of Leeds, Leeds LS2 9JT, SW7 2AZ, UK
[b] Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK
[c] Resources and Waste Advisory Group Ltd., Wren House, AL1 1NG, UK

HIGHLIGHTS

- Socioeconomic development indices can explain variability in cities' waste management performance.
- SDG11.6.1 indicator: controlled recovery and disposal: median ca 45 % for cities in low-income countries.
- Improvements in service quality often lag those in service coverage (i.e. extent).
- No overall evidence of decoupling socio-economic growth with waste generation.
- Machine learning and non-linear regression models each provide different insights

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Frederic Coulon

Keywords:
Municipal solid waste
Development indices
Sustainable development goals
Circular economy
Random forest

ABSTRACT

Mismanaged municipal solid waste (MSW), the major source of plastics pollution and a key contributor to climate forcing, in Global South cities poses public health and environmental problems. This study analyses the first consistent and quality assured dataset available for cities distributed worldwide, featuring a comprehensive set of solid waste management performance indicators (Wasteaware Cities Benchmark Indicators – WABI). Machine learning (multivariate random forest) and univariate non-linear regression are applied, identifying best-fit converging models for a broad range of explanatory socioeconomic variables. These proxies describe in a variety of ways generic levels of progress, such as Gross Domestic Product – Purchasing Power per capita, Social Progress Index (SPI) and Corruption Perceptions Index. Specifically, the research tests and quantitatively confirms a long-standing, yet unverified, hypothesis: that variability in cities' performance on MSW can be accounted for by socioeconomic development indices. The results provide a baseline for measuring progress as cities report MSW performance for the sustainable development goal SDG11.6.1 indicator: median rates of controlled recovery and disposal are approximately at 45 % for cities in low-income countries, 75 % in lower-middle, and 100 % for both upper-middle and high-income. Casting light on aspects beyond the SDG metric, on the quality of MSW-related services, show that improvements in service quality often lag improvements in service coverage. Overall, the findings suggest that progress in collection coverage, and controlled recovery and disposal has already taken place in low- and middle-income cities. However, if cities aspire to perform

Abbreviations: AICc, Akaike information criteria corrected; CPI, Corruption perceptions index; EPI, Environmental protection index; ESM, Environmentally sound management; EU, European Union; GDP, Gross domestic product; GIZ, Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH; GNI, Gross national income; H, High income - development group; HDI, Human development index; L, Low income - development group; L-M, Lower-Middle income development group; MSW, Municipal solid waste; OECD, Organisation for Economic Co-operation and Development; PPP, Purchasing power parity; R, R programming language; SDG, Sustainable development goal; SMAPE, Symmetric mean absolute percentage error; SPI, Social progress index; SWM, Solid waste management; U-M, Upper-Middle income - development group; UN, United Nations; WaCT, Waste Wise Cities Tool; WABI, Wasteaware Cities Benchmark Indicators; WaW2.0, World Bank Group 'What a Waste 2.0'; WBG, World Bank Group.

* Corresponding author.
E-mail address: c.velis@leeds.ac.uk (C.A. Velis).

http://dx.doi.org/10.1016/j.scitotenv.2023.161913
Received 19 October 2022; Received in revised form 25 January 2023; Accepted 26 January 2023
Available online 11 February 2023
0048-9697/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

better on MSW management than would have been anticipated by the average socioeconomic development in their country, they should identify ways to overcome systemic underlying failures associated with that socioeconomic level. Most alarmingly, 'business as usual' development would substantially increase their waste generation per capita unless new policies are found to promote decoupling.

1. Introduction

The concept of 'waste', its management to protect public health and the environment and its reduction through the circular economy, is key to sustainable development. Solid waste collection was introduced as part of the public health response to cholera epidemics (Strasser, 1999), and environmental controls explicitly associated with waste recovery and disposal were introduced from the 1970s (Wilson, 2007), yet substantial amounts of mismanaged MSW remain, particularly in the low- and middle-income countries that are still perceived as environmentally developing (often collectively referred to as 'Global South') (UNEP and ISWA, 2015). Mismanaged MSW is the major source of plastics entering the oceans (Jambeck et al., 2015; Lau et al., 2020), and contributes substantially to climate change via biodegradation of food wastes generating methane (Bogner et al., 2008) and open burning of plastics generating black carbon (Wiedinmyer et al., 2014; Reyna-Bensusan et al., 2019).

To progress further towards sustainable waste and resource management, and wider circular economy, in cities, requires an understanding of the current baseline and of how a city's performance varies with its socioeconomic conditions. Previous global comparisons have been limited and constrained by scarce and inconsistent data, due to a lack of standard definitions, measurements, and standard methodologies and systems for reporting (UNEP and ISWA, 2015, Kaza et al., 2018a, 2018b). Key recent reports by the World Bank Group 'What a Waste 2.0' (WaW2.0, hereafter) (Kaza et al., 2018a, 2018b) and a related update (Kaza et al., 2021), compiled available country and city data, featured curve-fitting only for waste generation and associated with just country level to Gross Domestic Product – Purchasing Power Parity (GDP-PPP) and it remains unclear whether point by point quality control was applied. Significant recent general worldwide overview analyses (Kawai and Tasaki, 2016; Bundhoo, 2018; Das et al., 2019) do not include new statistical analysis; only one used econometrics to relate waste treatment/disposal at country level to GDP-PPP (Tisserant et al., 2017). A still untested theoretical proposal 20 years ago suggested that solid waste management could be a useful and visible proxy indicator of urban governance (Whiteman et al., 2001). An empirically informed theorising on stages of progression of waste management systems in association with developmental level was published recently but features no quantitative data analytics (Whiteman et al., 2021). Much-needed future scenarios studies (Gómez-Sanabria et al., 2022) therefore are inevitably based largely on questionable baselines and speculated or indirect associations with level of socioeconomic development.

In this original data analytics a database is used (VelisEtAl2023_WABI_Input.xlsx), generated specifically to allow the comparison of cities' Solid Waste Management (SWM) systems on a consistent basis. This city profiling methodology was developed for UN-Habitat (Scheinberg et al., 2010); the data for the original 20 cities was later analysed with very limited curve-fitting against Gross National Income (GNI) per capita and Human Development Index (HDI) (Wilson et al., 2012). The methodology was further developed to become the Wasteaware Benchmark Indicators (WABI) (Wilson et al., 2015a), explained also in a detailed user manual (Wilson et al., 2015b). WABI include a comprehensive set of quantitative and qualitative indicators, covering the 'physical' and 'governance' aspects of integrated sustainable waste management, and organised within six different topic categories. They measure the performance of waste and resources management across global cities.

This study uses a sub-set of the set of quantitative WABI that are linked to the sustainable development goal (SDG) indicator 11.6.1 within SDG11 on sustainable cities (United Nations (UN), 2015) (Section 2.1: Data sources), which relates to the proportion of MSW collected and managed as a fraction of total MSW by cities (Fig. 1). Collection and disposal are assessed by traditional extent of service indicators coupled with novel composite indicators attempting to assess the quality of the services provided. In this context, the WABI database for 40 cities, the only current source of data on 'controlled recovery and disposal', with the data moderated to ensure comparability, is used to deliver a baseline for analysing the performance of cities. In the coming years, results will be made available for a new methodology called Waste Wise Cities Tool (WaCT), which is inspired in part by the WABI, and intends to provide a standardised methodology for data gathering just on SDG11.6.1 (UN-Habitat, 2020) (https://unhabitat.org/waste-wise-cities). Beyond the narrow reporting for the SDGs, the service level indicators corresponding to the SDG 11.6.1 are complemented here by parallel indicators of the quality-of-service provided, allowing deeper comparative insights between cities. These advantages outweigh the relatively small sample size of cities available to analyse – a side effect of the data-intensive nature of the WABI composite indicators. The WABI also include indicators for the 3Rs (reduction, reuse and recycling). These are clearly important for the overall performance of a city's combined solid waste and resource management system, but this paper focuses on aspects linked to SDG target 11.6.

This paper reports the first comprehensive quantified analysis to test the hypothesis *that variability in cities' performance on MSW can be accounted for by indices of socioeconomic development*, by applying a variety of regression analyses and machine learning models to a broad range of explanatory variables. The results offer statistically evidenced insights to permit debate on the performance of world cities in managing their waste, and why some cities appear to 'over-' 'or 'under-' perform compared to their socioeconomic development level.

2. Methods

2.1. Data sources

We collected and analysed data for the Wasteaware Benchmark Indicators (WABI) cities. The WABI methodology is a comprehensive set of indicators, aimed at measuring the overall performance of a city's municipal solid waste management (MSW) system (Wilson et al., 2015a). The WABIs methodological approach is available in a detailed manual (Wilson et al., 2015b) and is summarised here: WABIs overall methodological approach is based on a 'city profiling' process where the city's solid waste and resources management performance is assessed for a reference year by collation and analysis of mainly secondary data, complemented with on-site investigations. The assessment is performed for a series of indicators, many of which are composite. Those composite indicators in WABIs were generated by assessing against five or six criteria, with scores against each assessed at one of five levels, corresponding to zero, low, medium, medium/high or high level of compliance.

2.1.1. WABIs sub-set

We limited the analysis to the sub-set of WABI indicators relevant to the SDG 11.6.1 ('Waste Wise Cities' methodology, developed by the UN-Habitat as the custodian of SDG 11.6.1 (UN-Habitat, n.d.), defined as: '*Proportion of municipal solid waste collected and managed in controlled facilities, out of total municipal solid waste generated, by cities*'). The WABI approach can be considered as one of the main formative key bases for informing the recently released WaCT methodology, and are therefore directly relevant. The five relevant indicators are: (1) *Waste Collection*

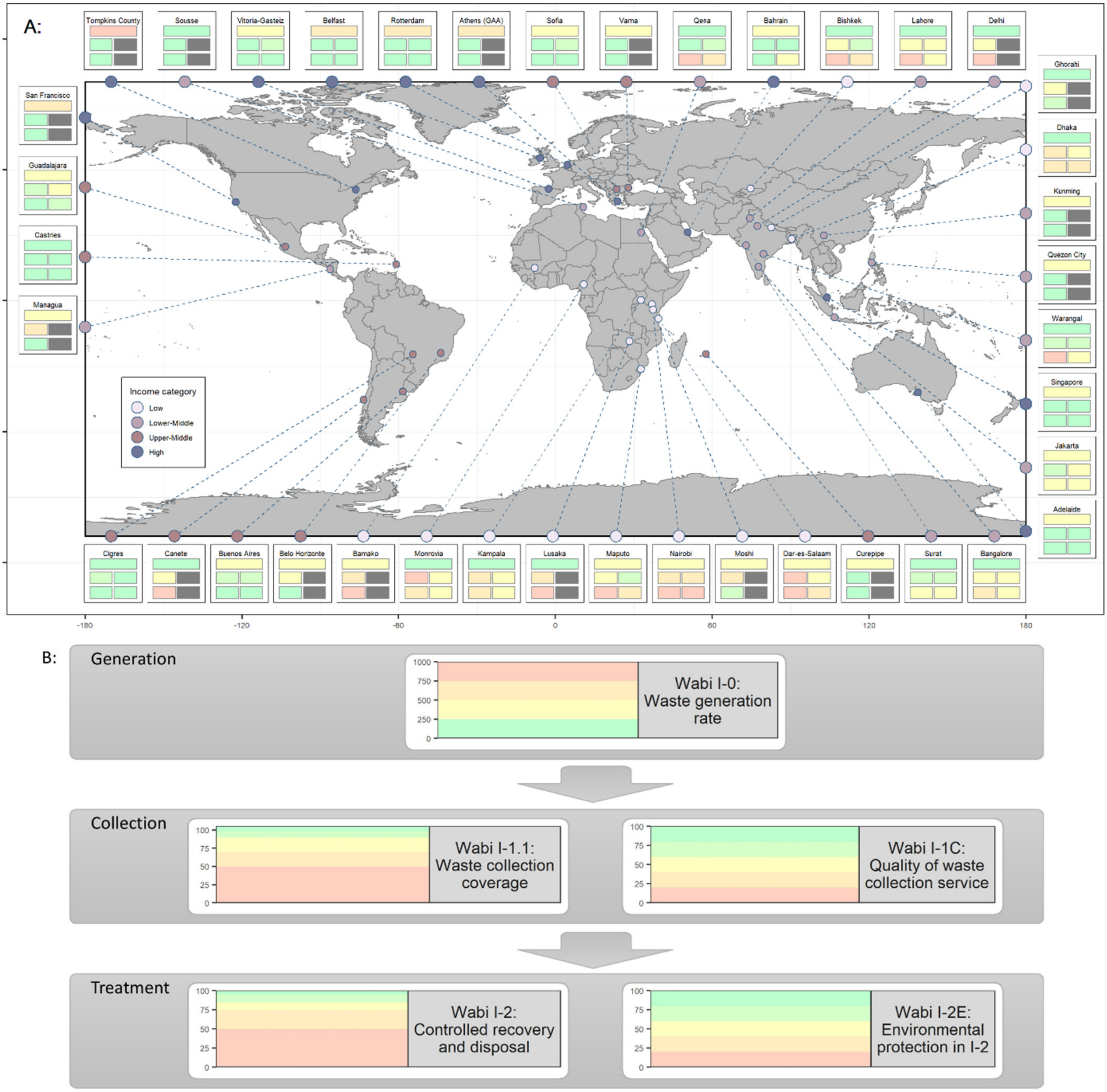


Fig. 1. A, B. (A): 40 cities across the World, covering all World Bank country income categories, for which their waste and resources management performance was consistently assessed with the Wasteaware Benchmark Indicators (WABI) methodology (Wilson et al., 2015a; Wilson et al., 2015b). Performance for five WABI indicators is stated for each by colour coding, from most well performing (green) to least well performing (pink), in line with the WABI methodology bands, banding which are also detailed in Figs. 3, 4 and 5. Dark grey denotes missing data. Colour coded country income categories showing cities in Low (L), Lower-Middle (L-M), Upper-Middle (U-M) and High (H) countries, as per the World Bank Atlas methodology (Section 2.1.2: WABI city cases dataset). (B): The 5 WABI used in this study – they all relate to aspects of the SDG indicator 11.6.1, defined as: '*Proportion of municipal solid waste collected and managed in controlled facilities, out of total municipal solid waste generated, by cities.*'

Coverage indicator (I-1.1), measuring directly or via a proxy the proportion of households served with a regular and reliable collection service; (2) *Quality of Waste Collection Service* (I-1C), which assesses the quality of the waste collection service against a set of six criteria using a standardised protocol; (3) *Controlled recovery and disposal* (I-2) indicator that measures the proportion of waste collected that is received at controlled recovery and disposal facilities (excluding what is reused or recycled); (4) *Degree of environmental protection in controlled recovery and disposal* (I-2E), the complementary to I-2, 'quality' indicator. We also included accompanying data on MSW generation, reported as part of the same city profiling efforts: (5) *Waste generation rate* (I-0), reported as total MSW generated per year per person (Kg.y^{-1}.p^{-1}), which is part of the background data collected, checked and reported when implementing the WABIs methodology.

2.1.2. WABI city cases dataset

The data were collected through the WABI profiling of 40 cities over a period of years (2009–2016) (VelisEtAl2023_WABI_Input.xlsx) facilitated by co-authors' custodianship (D.C.W, C.A.V and A.D.W.). Therefore, some of the cities were profiled with a previous version of the WABI, which did not include information on I-1C and I-2E: for these two indicators a smaller dataset was available. Selection of cities for profiling was random and ad-hoc, not identified with an up-front prescribed stratified sampling plan, but informed by the willingness of the city authorities and the availability of resources to conduct the WABI profiling. Despite this potential methodological limitation, the city profiling resources and quality assurance efforts were allocated so that the cities sample accounts for major variabilities encountered on the ground. First, dispersion of the case cities geographically covers the entire world (six inhabited continents) (Fig. 1), encompassing cultural variability. Second, they are spread across the entire socioeconomic development spectrum, as estimated via the GNI per capita, as recommended by the World Bank by categorising countries into four main development groups using their World Bank Atlas methodology (https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries): Low (L), Lower-Middle (L-M), Upper-Middle (U-M) and High (H). Third, the case cities also are spread across different population sizes: from 27 thousand, Ghorahi, Nepal, to 16 million for Delhi, India.

2.1.3. Explanatory variables

Our hypothesis is that the variability in cities' performance on MSW can be accounted for by indices of socioeconomic development. We have tested this hypothesis by applying a variety of regression analyses and machine learning models (Section 2.3: Data analysis – models and validation) to a wide set of explanatory variables (independent predictors), each of them broadly associated with level of socioeconomic development of a country. An initial set of 30 potentially suitable explanatory variables was collated based on the authors' expert opinion and by reference to the limited relevant literature and preliminary runs of models led to progressive elimination for the two-thirds of them down to 9 (Table S4 and Fig. 2) due to data availability or because of not rendering any statistically converging models. The key inclusion criterion for selecting the final set to run the complete data analytics was the comparatively wide availability of these explanatory variables for cities around the world: this could theoretically enable future users of our results to interpolate for cities that were not profiled by WABI, subject to suitable models been fit. Notably, such a final set of explanatory variables had to be decided, because part of the analyses conducted (Section 2.3.2: Approach 2: conditional random-forest) accounts for potential 'interactions' between them.

The country for each city and reference year in which it was profiled was used to obtain the corresponding value for each of the 9 independent explanatory variables, because these change over time. Notably, all 9 explanatory variables assess performance across countries rather than of individual cities. This is due to the absence of any comprehensive set of socioeconomic indicators describing the performance of cities across the world. A possible alternative could have been use of the gridded city/ regional GDP-PPP and HDI values (Kummu et al., 2018); however, these are available only for the biggest metropolitan agglomerations of the world and this would force us to limit our analysis only to those, and would render results not relevant for smaller cities in our dataset. Furthermore, not even this indicator is measured at local level, but estimated by an allocation process from regional level data, introducing therefore different additional sources of uncertainty. In any case, the use of country-level indicators here results in introducing a possible bias in each of the independent explanatory variables, to the degree that the performance of each city deviates from the country average. Therefore, it renders fitting converging models and the examination of our hypothesis more difficult. We speculate that if we were able to use socioeconomic indicators at city level, it would have been easier to discover potential associations and the prediction intervals should in principle decrease; yet this is still to be tested and

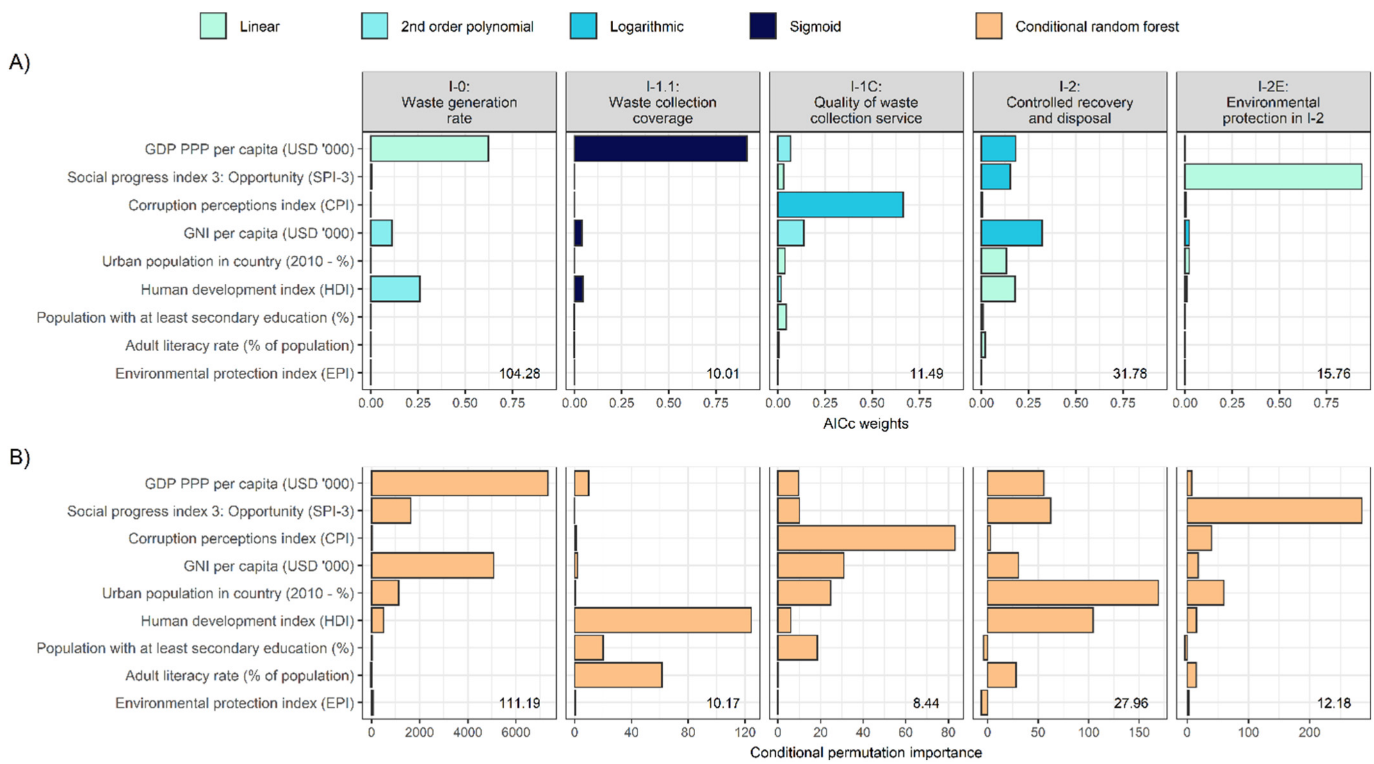


Fig. 2. A, B. Model selection analysis for each WABI indicator. (A): Best-fit regression model (Section 2.3.1: Approach 1: non-linear regression analysis): GDP PPP related variables were the most important for the waste generation rate and waste collection coverage, and variables relating to social progress were the most important for the controlled recovery and disposal aspect. Less dominance by a single variable was observed for the quality of waste collection services. Bar colour indicates the regression formula. (B): Multivariate conditional random forest analysis (Section 2.3.2: Approach 2: conditional random-forest): Relative variable importance is listed for the predictive equations for best explanatory variable. Comparative: RMSE values are given in the lower corner of each panel. I-1c and I-2E: There is agreement between the two approaches, but the multivariate approach predicts slightly better. I-0: There is an agreement, but univariate is slightly better than multivariate. GDP can be used as a good predictor by itself. I-1.1: Very strong unimodal dominance, and no agreement between the approaches. Almost no difference in predictive ability. I-2: No dominance in the univariate analysis, with a much better multivariate predictive performance.

should be considered in the future, if suitable datasets are made available in the literature.

2.2. Data preparation

2.2.1. WABI dataset quality assurance

The WABI dataset from published sources and background database (Table S2) for the five indicators considered here (Table S1 and as detailed elsewhere (Wilson et al., 2015a, Wilson et al., 2015b)), were cross-checked and corrected if needed. The WABI profile of each city has been quality assured via a triple stage process. First, the city profilers were selected to possess the waste and resource management expertise to understand relevant systems on the ground and trained on the WABI methodology; they are typically consultant/practitioners or supervised researchers, and have local knowledge. Second, the profiling outcome has been cross-checked by a group of experts, comprising the authors of the methodology for quality assurance purposes. Acceptable source data pedigree and scoring justifications were verified. Typical checks included confirming that the numbers reported resulted from accurate sourcing and were properly referenced, and the indicator definitions and guidance were accurately applied by the city profilers. For example, recurrent errors were identified, such as reporting waste generation rates that included construction and demolition waste. This step included checking the relevance for some indicators. For example, inclusion of information on informal recycling sector is anticipated for cities in the Global South. Third, accuracy of the composite indicator calculations was confirmed (in standard profiling process are largely automated via a spreadsheet). So each of the WABI data points is based on sourcing and collation of extensive secondary and primary data, rigorous processing and scrutiny for quality assurance effort to be collated.

Post the triple quality assurance, descriptive statistics was performed on each of the five dependent WABIs variables (WABI: I-0, I-1.1, I-1C, I-2, and I-2E) summarised in Table S3. Any apparent outliers/extreme values were examined for correctness. In very few certain occasions, this check revealed a difficulty to accurately define the value, due to ambiguity in the WABI definition or because of the special character of the city case. For example, Ghorahi sits alongside Maputo, as a *Low-income* category city achieving a comparatively high collection coverage. Here, the issue is where does one set the 'city' boundary: if one considers just the 'urban area', then the collection coverage is at 88 %, used here; if one considers the whole administrative area, including the unserved rural periphery, the collection coverage would drop to 52 %, closer to the current average line of Fig. 4A. This case also demonstrates the importance of a case by case quality management, as it was applied here.

2.2.2. Apparent outlier handling

No apparent statistical outlier/extreme value that remained after the quality control was excluded from the data analytics, despite the potential difficulty that could arise in fitting converging models. These were assumed to be strong yet actual over- or under-performing cases, from which valuable insights could be gained. Where feasible, explanations were sought in the contextual information about the waste and resources management in that city either in the detailed WABI city profile justification or in the literature. Any gaps in data (dependent and independent variables) were suitably coded so that they do not affect the data analytics.

2.3. Data analysis – models and validation

Data analytics and plotting were performed in R (4.0.3) (https://www.r-project.org/), including generic descriptive statistics, for example for the statistics displayed the box-plots. To test our overarching hypothesis, we used two complimentary approaches: (1) non-linear regression analysis, and (2) conditional random-forest.

2.3.1. Approach 1: non-linear regression analysis

We relied on non-linear regression analysis under a model selection approach to understand which explanatory variable best explains each dependent variable, while allowing several potential forms of the relation. As we were mostly interested in comparing the different explanatory variables relation to each of the different WABI and less in generalization to additional cities; and given the size of the dataset, we have not split the data to separate training and validation sets. Instead, we repeated the following procedure for each of the five dependent variables. First, we removed from the dataset any case with missing values for the dependent or any of the 9 explanatory variables. Next, we fitted for each pair of dependent and explanatory variable four models, using either a linear, a second order polynomial, a logarithmic or a sigmoid formulas (Table S5 and VelisEtAl2023_WABI_Summary.xlsx). Non-linear regression analysis was based on the Levenberg-Marquardt algorithm for model convergence (package 'minpack.lm' in R). We assessed the performance of each fitted model with both the *Root mean square error* (RMSE, Eq. (1)) and the *Symmetric mean absolute percentage error* (SMAPE, Eq. (2)), as well as *Akaike Information Criteria Corrected* for small sample size (AICc).

$$RMSE = \sqrt[2]{\left[\Sigma_n^N (\hat{y}_n - y_n)^2 / N\right]} \tag{1}$$

$$SMAPE = 2 \cdot \left[\Sigma_n^N |\hat{y}_n - y_n| / (|\hat{y}_n| + |y_n|)\right] / N \tag{2}$$

with N being the number of cases, $\hat{y}_n$ the predicted value for case n and y_n the observed value of case n. We also estimated the confidence intervals at significance level $\alpha = 0.05$ (the range of a mean value for a given X: 95 % that the true mean value falls within this intervals); and prediction intervals (range of a single observation for a given X: 95 % that a new sample would fall within these interval). We relied on the second order polynomial Taylor expansion as implemented in the 'propagate' R package (propagate::predictNLS).

After fitting a total of 36 models (4 shape functions and 9 explanatory variables) per dependent WABI variable, we performed a two-step model selection analysis using the calculated AICc values. In the first step, we identified for each explanatory variable the formula that had the lowest AICc value (out of the four fitted model). We then kept only the single best model of each explanatory variable and used it to identify the overall best model and to calculate the AICc weight of each explanatory variable. We have not taken all 36 models in a single model selection table, since some of the formulas can converge to one another (e.g., second order polynomial can become a linear model if the a_3 coefficient converges to 0), thereby, artificially inflating the relative importance of the explanatory variable. The two-step model selection procedure allowed exploration of multiple potential shapes of relation for each explanatory variable, while comparing the explanatory variables with equal representation.

The advantage of the non-linear regression analysis elaborated here is that projections can be made with a single variable, increasing therefore the potential coverage to many city cases for which we do not have detailed data on the socioeconomic indices (explanatory variables). We note, however, that our dataset coverage is limited, and that we have not explored quantitatively the models' ability to generalize to non-sampled cities. Nonetheless, basing the analysis on a single variable is considerably less demanding on information availability for explanatory variables, e.g. possible data gaps in HDI or any other socioeconomic index in comparison to 'Approach 2' where multiple explanatory variables are required.

2.3.2. Approach 2: conditional random-forest

A key limitation of the simple non-linear regression approach implemented here (Approach 1) is that it does not account for potential interactions between explanatory variables. Adding interaction to the candidate model list would increase its length considerably, especially if we wished to explore various potential curve forms for each explanatory variable and interaction term, and/or allow complex interaction between more than two explanatory variables. Therefore, we used conditional random forest (Strobl et al., 2008; Hapfelmeier and Ulm, 2013) to explore the relative importance of the different explanatory variables, while allowing for complex interactions between them. For each dependent variable we fitted a conditional random-forest algorithm against all explanatory variables as implemented in the 'party' R package, using the maximum number of cases with full data. Each conditional random-forest was based on 1000 trees, with the minimum sum of weights in a node to be considered for splitting (minsplit) and the minimum sum of weights in a terminal node (minbucket) set to the number of cases divided by 4 and 6, respectively. We then extracted for each explanatory variable the conditional variable importance, based on the change in mean decrease in model accuracy when conditionally permutating each variable. We used the RMSE index to compare the predictive ability of conditional random forest model to the top ranked model (high highest AICc weight) for each dependent variable. As we were mostly interested in the comparative variable importance, which necessitates permutation of the same data that is used to train the model, we have not applied a cross-validation procedure for validation.

2.4. Comparison with WaW2.0 (waste collection coverage)

For the majority of the WABI indicators there are no published statistics with worldwide coverage of cities to compare against in the literature. The most relevant information pertains to the WABI *Waste Collection Coverage* (I-1.1) indicator for which some partly comparable data were released as part of the datasets of WaW2.0 publication of the World Bank Group (WBG) (Kaza et al., 2018a, 2018b). However, these are only presented in bar-charts of individual values per WBG income category (Figs. 3.4, 3.10, 3.16 and 3.21 in the WaW2.0 publication). To this we have performed basic descriptive statistics to the WBG dataset (Kaza et al., 2018a, 2018b). Before that the dataset was cross-checked for obvious mistakes/typos (impossibly extreme values: e.g. only 40 % collection coverage for Wellington, New Zealand, one of the most environmentally developed places around the world) and these were omitted from the analysis.

Caution should be exercised in the comparative assessment of these two datasets due to methodological differences in the definition and wider data processing. For the WABI definition, *Waste Collection Coverage* is measured directly or via a proxy as the proportion of households served with a regular and reliable collection service. For the WaW2.0 four different definitions are possible and the maximum (least conservative) value of the four possible ones, whichever are available is selected as the true value. Another methodological difference relies on the representativeness of the datasets: while the WaW2.0 initially contained 250 cities (a much higher number than the 40 in WABI), no quality control has been applied to cross validate the individual data entries resulting in certain erroneous values, and there is a potential positive bias towards Indian cities (39 in the database) in comparison with China (only one: Beijing); vs. WABI which are distributed all over the world. As explained, no data analytics (statistics or curve fitting) has been reported by the WBG.

3. Results

3.1. Global baseline links to socioeconomics

The WABI database points used in this work (Fig. 1) provide a worldwide baseline, mapping 40 cities, across the spectrum of population size, the four major country income categories and all six inhabited continents. For each city, data are shown for waste generation per capita, collection coverage and controlled recovery (treatment) and disposal. This analysis establishes that it is possible to use country-level socioeconomic development indices to explain the variability in performance of MSW management systems in cities around the world.

Fig. 2 summarises the results for 5 WABIs included in Fig. 1, with the best-fit regression model for each single explanatory variable (Fig. 2A) and conditional variable importance from a multivariate conditional random forest analysis (Fig. 2B) which takes as input all 9 explanatory variables and accounts for potential interactions between them (Section 2.3.2: Approach 2: conditional random-forest). GDP per capita (using the purchasing power parity PPP definition) is the most powerful single explanatory variable for two of the WABI, waste generation rate and collection coverage. Each of the other three WABI has a different dominant single variable, namely social progress index - opportunity (SPI-3: measuring 'opportunity' aspects), corruption perception index (CPI) and gross national income (GNI – Atlas definition) per capita. For the multivariate analysis, the best fit models involve significant contributions from between four and seven of the nine variables. Additional significant variables are human development index (HDI), urban population in a country, adult literacy rate and percent of population with at least secondary education. The only variable that was intuitively considered relevant, but which showed comparatively negligible explanatory power was the environmental protection index (EPI). Notably, the EPI version used here included measure of waste management (Table S4), which was dropped in the recent EPI definition updates but the dimension, informed by the WABI evidence presented here, should likely be reinstated in the future version.

Regarding predictive power, the random forest multivariate models generally outperform the univariate models. For service quality indicators, the multivariate root mean square error (RMSE) is 36 % lower than the best univariate model for collection, and 29 % lower for recovery and disposal; while for the controlled disposal rate, RMSE is 14 % lower. For collection coverage, the two approaches are similar: the sigmoid fit using GDP-PPP has RMSE 1 % lower than the multivariate model, although the latter makes minimal use of GDP-PPP. The exception is waste generation per capita: the single variable GDP-PPP linear regression has RMSE 7 % lower than the best multivariate model.

3.2. Socioeconomic development associated with a linear increase in MSW generation

Waste generation rate per capita is one of the most important and widely reported indicators for SWM systems, and is critical for formulating waste management plans, investment projects and operations. Expressed here it corresponds to the *I-0: Waste generation rate*, an addition to the standard WABI. The MSW data for cities show a wide range, from 119 to 995 $Kg.y^{-1}.p^{-1}$, with median at 276 $Kg.y^{-1}.p^{-1}$. *Low* and *Lower-middle* income countries show a similar median, within the 'green' band threshold of 250 $Kg.y^{-1}.p^{-1}$; which increases substantially with income level (Fig. 3D).

Several socioeconomic indices can serve as an explanatory variable capable of modelling the variability in MSW waste generation rate of cities (Fig. 2). The best fit is linear with country-level GDP-PPP per capita (Fig. 3A); followed by a second order polynomial relationship with HDI (Fig. 3B); and third best a country-level GNI per capita (Fig. 3C). Each of these univariate models has better predictive ability than the best fit random forest multivariate model (Fig. 2). There is considerable variability in the data, but a very strong positive correlation between the amount of waste generated in urban environments by individuals and the wider level of socioeconomic development of the country is observed. This confirms previous positive correlations reported in the GWMO (UNEP and ISWA, 2015) and WaW2.0 (Kaza et al., 2018a, 2018b). Most previous analyses have been made at country level, and have been constrained by poor and inconsistent data; indeed, researchers (Kawai and Tasaki, 2016) concluded that their compiled national dataset was too 'noisy' to fit a correlation with GDP. Predicted generation rates for cities from the model applied here are consistently higher than those predicted for entire countries by WaW2.0 (Kaza et al., 2018a, 2018b), which concurs with expectations that living standards, consumption and waste generation are higher in urban than in rural areas, particularly in low- and middle-income countries.

From all available data waste generation per capita is much greater in higher-income than lower-income countries. Historical data suggest substantial increases took place between 1980 and 2000 – with a 58 % increase in waste per capita observed for OECD data for the then EU-15 (The Association of cities and regions for recycling and sustainable resource management (ACR+), 2009), while GDP-PPP per capita over the same period grew by around 100 % (OECD, 2020). In the following five years, GDP increased by 70 % (USD $27,325 in 2005) (Macrotrends, 2020), while waste per capita only grew by 4.6 % (The Association of cities and regions for recycling and sustainable resource management (ACR+),

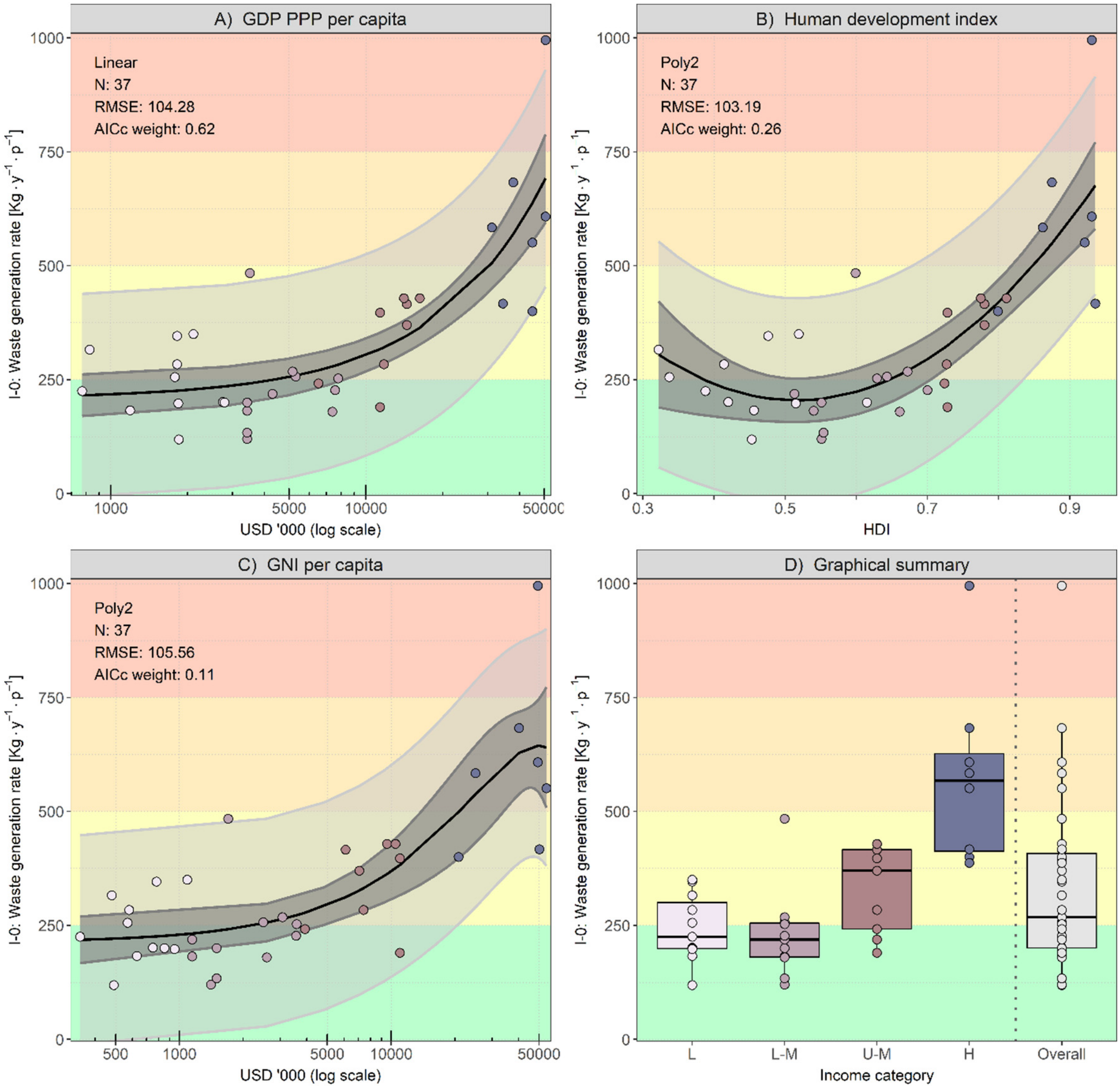


Fig. 3. A–D: Waste generation rate. (A): Waste generation rate is best explained by a linear model against the mean GDP-PPP per-capita associated with the country the city belongs to and for the year of the waste generation rate measurement. Black line is the fitted curve, dark grey shade the confidence interval and light grey shade the prediction intervals, while points colour corresponds to the income categories of (D). (B): Second-best fit is by country level human development index (HDI). (C): Third-best fit is by country-level Gross National Income (GNI) per capita (Atlas definition). (D): Summary statistics: overall and per GNI income category, as defined by the World Bank Group. Boxes represent the inter-quartile range with the median as a horizontal line. Hinges are 1.5 times the inter-quartile range, while points outside the hinge are statistical outliers. Colour coding in four performance bands: most well performing (green) to least well performing (pink), performance defined here for absolute quantities, irrespective of socioeconomic category.

2009), which led to multiple claims that 'decoupling' was taking place. Several studies concluded at the time that the evidence for such decoupling was at best weak (Mazzanti and Zoboli, 2008; Fell et al., 2010). Since then, several authors have searched for evidence at a relatively local level in for example Switzerland (Jaligot and Chenal, 2018) and Australia (Madden et al., 2019). The notion of 'decoupling' despite dominating the agenda for implementing the post-2015 SDGs, is loosely defined (Fletcher and Rammelt, 2017). Even if decoupling is eventually proven to exist at higher income levels, the strong link between waste generation and socio-economic development still exists at lower-income levels.

3.3. Assessment of quality of waste collection needed along with collection coverage

Collecting waste is a core utility service underpinning public health; is embedded within SDG targets 1.4 and 11.6, and is explicitly measured as a core part of SDG indicator 11.6.1: the proportion of total waste generated that is collected (UN-Habitat, 2020). The WABI dataset names this *Waste collection coverage* (WABI: I-1.1), which sits alongside a parallel indicator which assesses the *Quality of waste collection service* (WABI: I-1.C) against a set of six criteria using a standardised protocol (Wilson et al., 2015a). The results are summarised in Fig. 4 and descriptive statistics in Supplementary Information (SI) (SI.3, Table S3). Fig. 4A shows the best fit curve using a sigmoid model against GDP-PPP per-capita. The inflection point of the curve indicates a division into two main parts. At lower-income levels, collection coverage increases roughly linearly. At higher levels most cities approach or reach the SDG target of universal waste collection; this transition is estimated here as between USD 5000–8000 GDP-PPP per capita. Some lower-income countries, however, are observed to perform better than some high-income cities.

Assessing the performance of a city's waste collection requires indicators for both service level (collection coverage) and service quality (Fig. 4A and B, respectively). Data for the latter (WABI: I-1C) feature a minimum at 29 % and median at 66 %, unlike collection coverage, where the one peak is around 100 % (histogram at Fig. S1 – WABI: I-1.1). Fig. 2

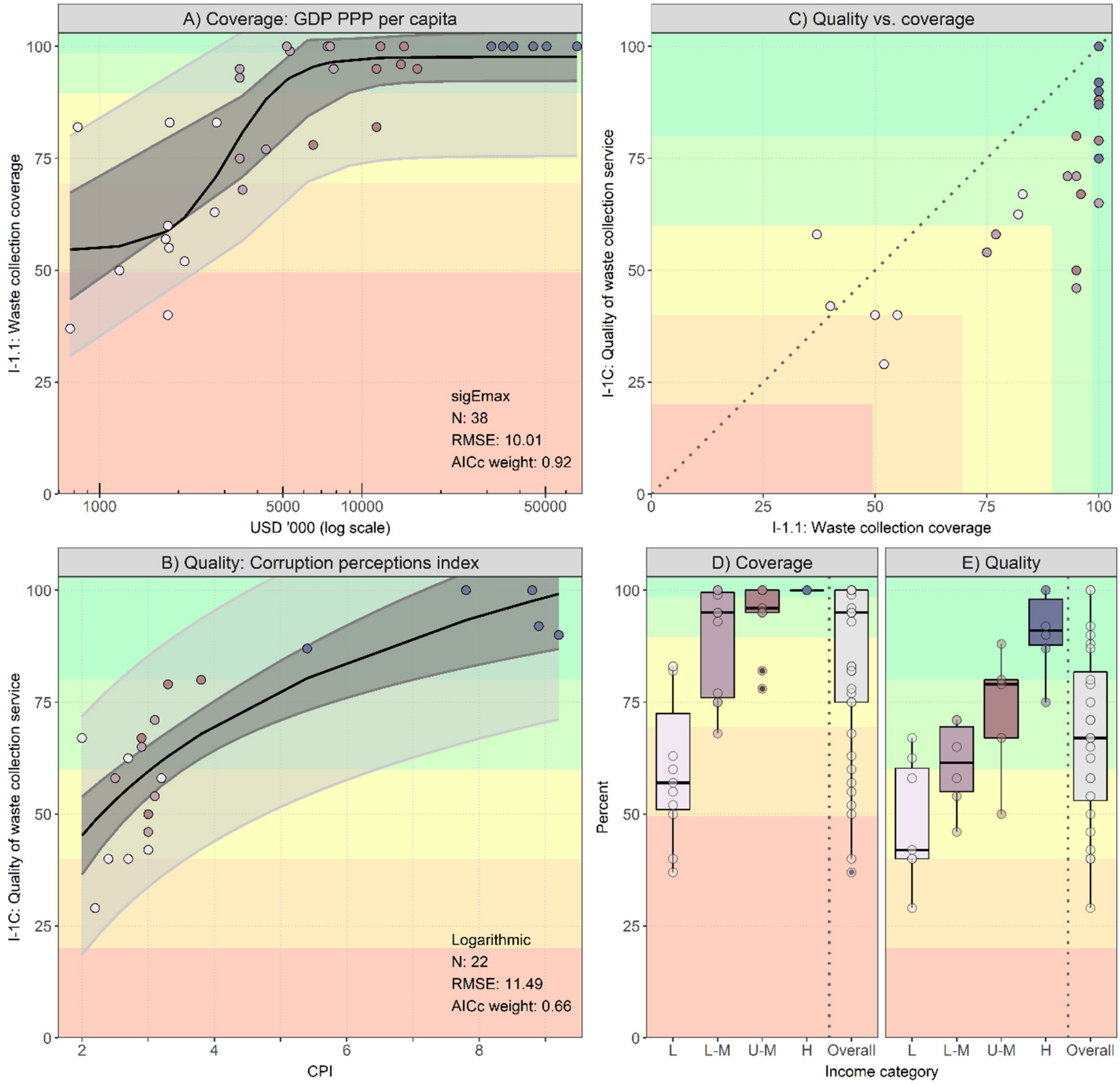


Fig. 4. A–E: Waste collection and quality. (A): Waste collection coverage was best explained by a sigmoid model against the GDP-PPP per-capita for the year in which the variable was measured. (B): Quality of waste collection coverage was best fitted by a logarithmic model against the *Corruption perceptions index*. (C): Plotting the quality of waste collection coverage against waste collection coverage reveals that most cases are below the line of unity (dotted black line). (D, E): Summary statistics for the two WABI variables: overall and per GNI income category, as defined by the World Bank Group. Boxes represent the inter-quartile range with the median as a horizontal line. Hinges are 1.5 times the inter-quartile range while points outside the hinge are statistical outliers. In (A) + (B), black lines are the fitted curve, dark grey shades are the confidence interval and light grey shades are the prediction intervals. Points colour corresponds to the income categories of (D) and (E). Colour coding in five performance bands: most well performing (green) to least well performing (pink), in line with the WABI methodology bands.

showed that a relatively complex 7-variable multivariate model provides the best fit, with an RMSE 36 % better than for the best fit regression model, which was a logarithmic model of *Corruption perceptions index* (Fig. 4B). Comparing the extent of coverage vs. quality in Fig. 4C most cities lie below the line of unity, showing that improvements in service quality often lag improvements in service coverage.

3.4. Progress through elimination of open dumping and burning

WABI indicator I-2 is *Controlled recovery and disposal*, that is the proportion MSW remaining after recycling and reuse which goes to either a state-of-the-art, engineered or 'controlled' recovery/disposal site (Table S1). Achieving 100 % on this indicator eliminates open dumping and burning of MSW; which aligns with the second part of SDG indicator 11.6.1 and an important step towards environmentally sound management (SDG target 12.4) (UNEP and ISWA, 2015). The definition of what is a 'controlled' facility focuses primarily on operational control, rather than engineering/design (Wilson et al., 2015a; UN-Habitat, 2020). Fig. 2 shows that a six variable model gives a 14 % better fit to the data than the best univariate model: the dominant explanatory variables are percentage of urban population in the country and HDI. Five different single variables show limited explanatory power: the best fit is logarithmic with GNI per capita (Fig. 5A). The graph reflects two spikes in the distribution (bimodal) of this indicator in the WABIs database (Fig. S1 – WABI: I-2), one at 100 % (20 cities) and the other at 0 % (7 cities) – each facility is either controlled or not, and many cities have relatively few recovery/disposal facilities, so that many cities may be scored as either entirely 'controlled' or entirely 'uncontrolled'. It should be noted that 'controlled' status is not an end in itself, but corresponds to 'basic control' on the recent WaCT five point 'control ladder' (none, limited, basic, improved, full control) (UN-Habitat,

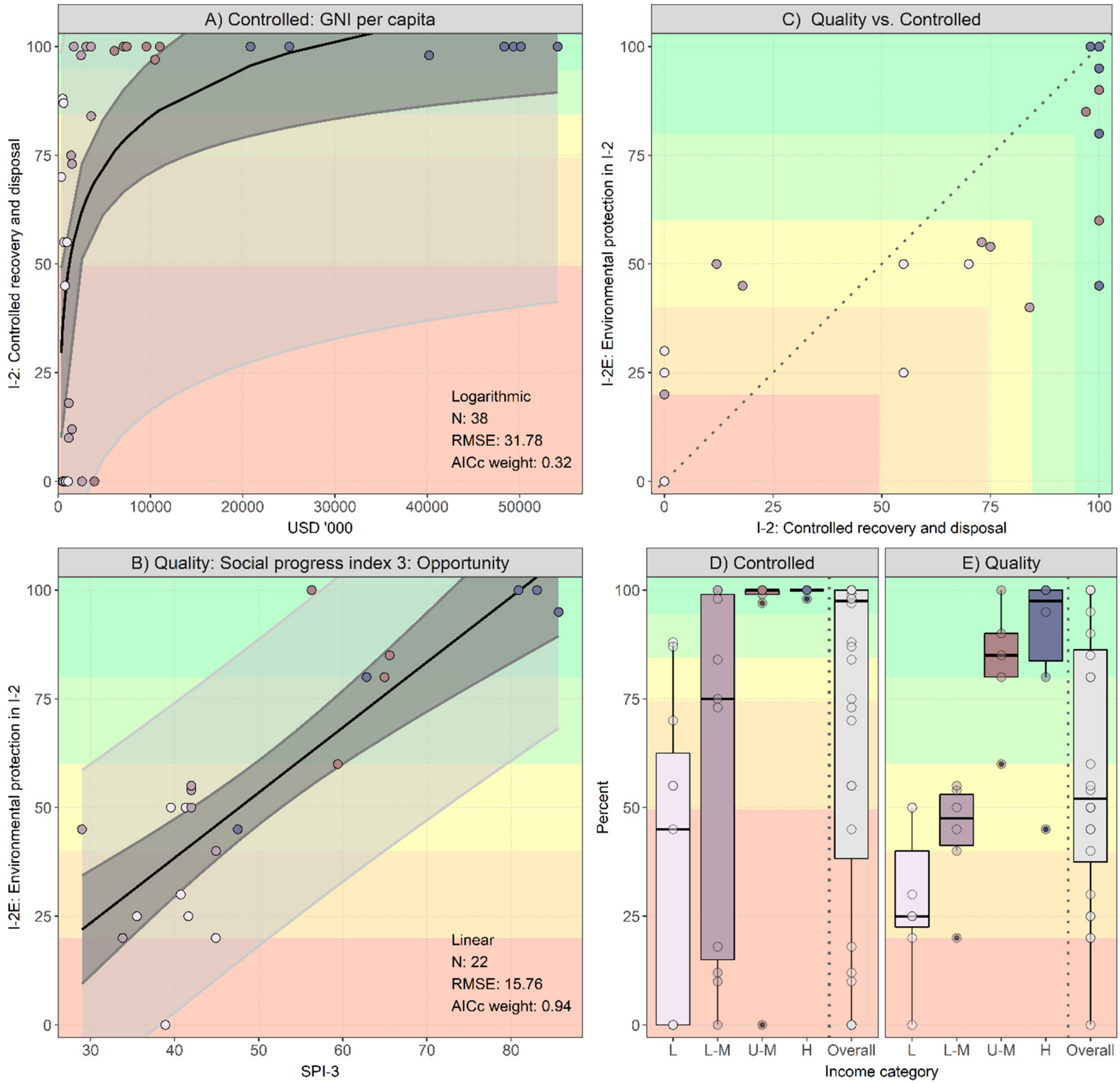


Fig. 5. A–E: Waste recovery and disposal. (A): WABI: I-2, the rate of controlled recovery and disposal was best explained by a logarithmic model against the social progress index. (B): The quality indicator WABI: I-2E *Environmental protection of waste recovery and disposal* was best explained by a linear model against the social progress index 3 (opportunity). (C): Plotting the environmental protection against the controlled recovery and disposal a trade-off between the two (the dotted black line is the line of unity). (D, E): Summary statistics for the two WABI indicators: overall and per GNI income category, as defined by the World Bank. Boxes represent the inter-quartile range with the median as a horizontal line. Hinges are 1.5 times the inter-quartile range while points outside the hinge are statistical outliers. In (A) + (B), black lines are the fitted curve, dark grey shades are the confidence interval and light grey shades are the prediction intervals. Points colour corresponds to the income categories of (D) and (E). Colour coding in five performance bands: most well performing (green) to least well performing (pink), in line with the WABI methodology bands.

2020), where full control relates to 'environmentally sound management' (ESM) as in SDG 12.4.

The WABI indicator for controlled recovery and disposal (I-2) is accompanied by a complementary indicator of service quality, here termed the degree of environmental protection (I-2E). Fig. 2 shows that the dominant explanatory variable for the quality indicator is the social progress index – opportunity 3 (SPI-3), which comprises sub-indicators on: *Personal Rights*; *Personal Freedom and Choice*; *Tolerance and Inclusion*; and *Access to Advanced Education.* Including relatively small contributions from six other variables enables the best fit multivariate model to provide a 30 % lower RMSE than the linear best fit univariate model (Fig. 5B). As in all the correlations, all the cities with available data have been included; including in this case, the explainable high-performing case of Sofia, Bulgaria, which meets EU standards for waste management. Fig. 5E shows that the median degree of environmental protection increases with income level category.

Comparing their rate of control with their *Degree of environmental protection* (WABI: I-2 vs. 1-2E) in Fig. 5C, some lower-income cities sit above the line of unity while most cities lie below. This can be explained as the WABI definition of a controlled facility requires at least a medium rating (10 out of 20) for three of the six 'quality' criteria, relating to the degree of control over *waste reception and general site management*, and *waste recovery and disposal*, and the *degree of monitoring and verification of environmental controls*. Lower-performing cities often achieve some level of control, even though the facilities remain 'uncontrolled', while cities first achieving 'control' have some way to go to reach a high rating on these criteria, to equate with meeting the SDG target 12.4 of environmentally sound management (ESM). Similarly, most low- and middle-income countries will achieve average scores against the remaining criteria relating to efficiency of energy generation and use (where applicable); technical competence in planning, management, and monitoring; and occupational health and safety.

4. Discussion

4.1. Socioeconomic indices can model waste management variability

Using the first consistent and quality-assured dataset available for global cities it is demonstrated that country-level socioeconomic indicators can model much of the variability in performance of cities' MSW systems. The plurality of methodological approaches (multivariate random forest model vs. best univariate non-linear regression models) enabled the converging best fit models to be found, within the constraints of the relatively small dataset. These models provide a strong evidence-base for initial benchmarking of SDG indicator 11.6.1, which should allow cities' MSW performance to be compared using consistent and reliable data.

Specifically, the waste generation rate models developed can be used, alongside forecasts of population increases and of migration from rural to urban areas (United Nations (UN), 2018), to predict future waste arisings in most cities around the world (Hoornweg et al., 2013). For cities in low- and lower-middle income countries in particular, any business-as-usual scenario that locks them in the pathway of average baseline performance documented here will inevitably mean that waste generation per capita will increase as income levels rise; given the forecasts for city population growth and urban migration, this means that many cities in Africa and Asia would double their total MSW over 20 years (UNEP and ISWA, 2015). To prevent that happening, new policies are urgently needed to promote the 'decoupling' of waste growth from economic growth and the transition to a circular economy in developing countries as well as in developed countries – a prospect that is far from well-defined and certain (Preston and Lehne, 2017; Velis, 2017).

High variability for the waste generation rate is evident at each level of socioeconomic development. Some of this variability can be attributed to the degree of departure from the hypothesis that city-wide development performance can be represented by the average socioeconomic development of a country – an inevitable compromise in the analytical power of our approach in the absence of better practicable alternatives (Section 2.1.3: Explanatory variables). A next level of analysis would be to identify suitable explanatory variables at the city rather than the country level, which would allow comparison between cities within the same country, and to further extend the analysis to within the cities themselves to identify discrepancies between service levels at the neighbourhood level.

These results suggest that developing country cities have made greater progress in improving collection coverage than is generally acknowledged. The WaW2.0 report (Kaza et al., 2018a, 2018b) still provides the most extensive current baseline to date: in Table 1, the WABI median values (Fig. 4D) are compared with the WaW2.0 reported data, expressed as national average figures including both urban and rural areas. The much higher values reported by the WABI, by comparison, is explained as the dataset is derived for cities. The WaW2.0 database does include city as well as national data, which does not appear to have been used in the published report. Performing descriptive statistics here, however, using the raw WaW2.0 city dataset shows (Table 1 - last column) the data appear broadly to confirm the WABI results.

These data show that cities can make much more progress in extending waste collection to all their citizens than would be 'normal' for their income level. An example is Maputo, the capital of low-income Mozambique, which is the positive 'outlier' in the top left corner of Fig. 4. The city and its international technical assistance partners focused on extending collection coverage across the city as a priority in the early 2000s (Stretz, 2012; Stretz, 2013), resulting in a comparatively high 83 % collection coverage. Notably, as apparent outliers were not excluded (Section 2.2.2: Apparent outlier handling) in this analysis, Maputo's influence largely explains the apparent 'flattening' of the fitted curve towards higher values in Fig. 4A at the lowest-income levels (GDP-PPP per capita: $826). It is therefore, not anticipated that such flattening would hold in general.

Regarding quality of collection services (WABI: I-1C – see SI.1, Table S1), this is based on six assessment aspects (sub-indicators (Wilson et al., 2015b)) covering: (i-ii) two that relate to the degree of littering (at collection points or on the streets), (iii) uncollected wastes in lower-income districts; (iv) controls over waste transport, (v) service planning/monitoring, and (vi) health and safety provisions for collection workers. Averaging a '*medium*' score against each criterion would give an overall 50 % collection services quality indicator, while an average of '*medium/high*' would give 75 %. So, for cities with *medium/high* or *high* collection coverage (>90 %), it is understandable that service quality is extremely variable, with a range from around 50 % to 100 %. Common weaknesses in middle-income countries include lack of street-cleaning outside the central business district and more prosperous areas, issues with uncollected wastes in informal settlements and low usage of personal protective clothing and equipment by collection workers. The inclusion of littering phenomena, prevalent in disorderly urban settings (Weaver, 2015), may mean that some cities in high-income countries do not achieve a perfect 100 % score for service quality. For waste collection, SDG indicator 11.6.1 focuses on improving collection coverage, extending the service to all citizens. Therefore, the baseline level in cities is better than generally acknowledged, and even when 100 % collection has been achieved, attention will likely be required to continue to improve the quality-of-service provision.

Table 1
Comparing baseline values for waste collection coverage: our results vs. WaW2.0.

Socioeconomic development level	Waste collection coverage				
	WABI city (Median)	WaW2.0 report			WaW2.0 database***
		National averages*	Urban**	Rural**	Cities
Low income	57 %	39 %	48 %	33 %	60.5 %
Lower-middle	95 %	51 %	71 %	33 %	86.8 %
Upper-middle	96 %	82 %	85 %	45 %	98.5 %
High income	100 %	96 %	100 %	98 %	100 %

*Fig. 2.10 in the WaW2.0 publication. **Fig. 2.11 in the WaW2.0 publication. WABI – Wasteaware benchmark indicators. WaW2.0 – What a Waste 2.0 (Kaza et al., 2018a, 2018b) shown here only for completeness of information: there are based on questionnaires for country-wide data and cannot be directly compared with data pertaining to city administration boundaries, with which we deal here. ***Median values for the WaW2.0 database, post quality assurance performed in this study. Country income levels according to year of profiling for WABI dataset, and to reference year 2011 for WaW2.0 dataset. Socioeconomic development spectrum, as estimated via the GNI per capita, as recommended by the World Bank Group by categorising countries in to four main development groups: Low (L), Lower-Middle (L-M), Upper-Middle (U-M) and High (H).

On average, the income level of a country has a marked impact on the rate of controlled recovery and disposal (Fig. 5D). The median rates of control are approximately at 45 % for cities in low-income countries, 75 % in lower-middle, and 100 % for both upper-middle and high-income. These stand in sharp contrast to the figures given for open dumping (uncontrolled disposal) in the WaW2.0: 93 % (i.e. 7 % controlled disposal) in low-income countries; 66 % (i.e. 34 % controlled) in lower-middle, and 30 % (i.e. 70 % controlled) in upper-middle income countries (Kaza et al., 2018a, 2018b). This clearly demonstrates that, although much remains to be done to harmonise the methodological approaches to assessing city collection, recovery and disposal systems performance, cities can and are making progress in eliminating uncontrolled disposal and open burning of MSW.

Using the degree of environmental protection alongside the rate of controlled recovery and disposal adds a useful dimension to the analysis, certainly for middle-income countries, which are still journeying towards full ESM. It may even be useful for some developed country cities which, despite a near universal claim to a 100 % controlled disposal (or even ESM) rate, still struggle with a small but significant level of fly-tipping/illegal waste dumping and uncollected littering 'leaking' from their waste management systems (Liu et al., 2017). Among the high-income cities in the WABI database, Belfast is unusual in admitting to an estimated 2 % 'leakage', as it finds illegal disposal practices difficult to eliminate (Wilson et al., 2015a).

4.2. Evidence-based/targeted improvements are possible across the world

The raw data, descriptive statistics and multiple modelling attempts presented here, for a set of comprehensive waste and resources management indicators, allows us to start making sense in a more fundamental way of how cities are managing their solid waste. They constitute a novel and comparatively robust evidence base that can be used to put individual city performance in national and international / geographical context. New city profiles by the WABI methodology can be compared against this benchmark and used to revise this global analysis and benchmark with the approach released here.

Most importantly, the results allow clear differentiation on the potential priorities for cities around the world. Many cities at the higher end of the development spectrum have already achieved near universal collection coverage and ESM of the collected waste, but issues often persist around littering/fly-tipping and the effectiveness of street sweeping. Focussing efforts here could make a substantial difference to overall performance, but in many contexts performance aspect is not regularly monitored, even in the most affluent and data-driven cities. The analysis here demonstrates the value of collecting and analysing data in a coherent and robust way, and utilising such data in performance monitoring and management. In addition, such improvements could considerably abate their small but non-negligible contribution to the emerging global challenge of plastic pollution.

For cities towards the middle and lower developed end of the spectrum, the results demonstrate both the challenge and the opportunity. The challenge is that overall improvement in solid waste management performance is correlated with socioeconomic level, and rates of development can be slow, particularly outside of the most affluent / capital cities in each country. The opportunity is that the detailed city data analysed here clearly show substantial progress towards universal collection coverage (Indicator I-1.1, Fig. 4A) and towards controlled recovery and disposal (Indicator I-1, Fig. 5A) (i.e. high performance on indicators for SDG 11.6.1) even in some low-income cities, with specific cases that outperform considerably the average at their socioeconomic level. Therefore, major improvements in performance towards indicator SDG 11.6.1 are achievable – we speculate, given local political will and effective development support. Cities can buck the trend by seeking inspiration, lessons learned and transferrable best practices from the outperformers documented here and beyond (Scheinberg et al., 2010; Whiteman et al., 2021), especially from those with similar or slightly higher socioeconomic development indices to their own.

5. Conclusions and outlook

5.1. Casting light on SDG11.6.1

It is argued here that the WABI not only underpin, in part, the WaCT that measures the SDG11.6.1 indicator, but also substantially compliment the SDG assessment, by providing performance quality counterparts (quality-of-service) to the standard service level indicators of waste generation per capita, collection coverage and controlled recovery and disposal. Therefore, it offers additional insights, especially when a city achieves the initial compliance target. Such complementarity could in principle be further enhanced by increasing the number of cities in the database to improve the model fits; identifying suitable explanatory variables at the city rather than the country level, which would also allow comparison between cities within the same country; and extending the data analytics to the recycling and governance aspects of a city's MSW performance, factors also covered by the WABI.

5.2. A richer picture beyond SDG monitoring

This study demonstrates quantitatively that the global challenge of waste management is inherently linked to a city's socioeconomic development. If cities, however, aspire to perform better on MSW management than would be anticipated by the average socioeconomic development in their country, they should identify ways to overcome systemic underlying failures associated with that socioeconomic level which, in turn, can lead to the unlocking of development potentials across other urban management sectors. The variability encountered between cities at comparable level of socioeconomic development demonstrates that this, in principle, is feasible. Demonstrating the correlation between socioeconomic indices and SWM performance in turn adds further weight to the theoretical proposals that SWM could serve as a useful proxy of urban governance (Whiteman et al., 2001).

The models fitted here can form a basis for future predictions based on forecasting scenarios/modelling of socioeconomic development. Progress in collection coverage and controlled recovery and disposal is shown to be possible, and already taking place in low- and middle-income cities, but 'business as usual' development will continue to increase waste generation per capita unless new locally relevant policies can be found to promote decoupling and the transition to a more circular economy.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.161913.

Funding

C.A.V.'s research at the University of Leeds (UoL) was supported by GIZ grant 'Solid waste flow diagram tool and ISWM benchmark indicator monitoring for cities'. The relevant WABI web-portal was funded by EPSRC impact acceleration fund for Univeristy of Leeds (UoL internal allocation, via the 'Cities' cross-university theme') and Wasteaware Ltd. in-kind contribution.

CRediT authorship contribution statement

Costas A. Velis: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **David C. Wilson:** Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yoni Gavish:** Data curation, Formal analysis, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Sue M. Grimes:** Formal analysis, Project administration, Supervision, Validation, Writing – review & editing. **Andrew Whiteman:** Data curation, Funding acquisition, Validation, Writing – review & editing.

Data and code availability

All datasets used in this study are freely available online in Zenodo: Input (VelisEtAl2023_WABI_Input.xlsx), output (VelisEtAl2023_WABI_Summary.xlsx). The R code is also provided along with the 'Analysis Protocol – VelisEtAl Solid Waste Management Performance Cities – WABI – 2023'. https://zenodo.org/record/7585229#.Y9fWlRfP03s – DOI: 10.5281/zenodo.7585229.

Declaration of competing interest

Dr. Costas Velis reports a relationship with International Solid Waste Association (ISWA) that includes: consulting or advisory, funding grants, speaking and lecture fees, and travel reimbursement. Dr. Costas Velis and Prof. David C Wislon, Prof. Sue M Grimes are actively involved in the waste, resources and circular economy sector in research, advising, scholarship and consultancy capacities. Mr. Andrew Whiteman is regularly providing consultancy services for the sector worldwide.

Acknowledgements

We are grateful to UN-Habitat for funding the work on the initial version of the indicators, including 20 city profiles, and GIZ through their 'Operator Models' project which funded an intermediate version of the indicators and a further 5 city profiles. We acknowledge all the profilers of the individual cities – many are named in Table S2, and others in reference (Scheinberg et al., 2010). We thank past MSc students under the authors' supervision for offering preliminary partial data clearing and commentary: Henry Hickman (MSc dissertation at University of Leeds, supervised by C.A.V and D.C.W.) and Margaux Fargier (Final year MEng dissertation at Imperial College London, supervised by S.M.G, D.C.W and C.A.V.). We are grateful to Dr. Josh Cottom and Mr. Ed Cook at the University of Leeds for input on the GDP version selection. We acknowledge the support of Dr. Ljiljana Rodic for contributing in data quality control.

References

Bogner, J., Pipatti, R., Hashimoto, S., Diaz, C., Mareckova, K., Diaz, L., Kjeldsen, P., Monni, S., Faaij, A., Qingxian, G., Tianzhu, Z., Mohammed, A.A., Sutamihardja, R.T.M., Gregory, R., 2008. Mitigation of global greenhouse gas emissions from waste: conclusions and strategies from the intergovernmental panel on climate change (IPCC) fourth assessment report. Working group III (Mitigation). Waste Manag. Res. 26 (1), 11–32.

Bundhoo, Z.M.A., 2018. Solid waste management in least developed countries: current status and challenges faced. J. Mater. Cycl. Waste Manag. 20 (3), 1867–1877.

Das, S., Lee, S.H., Kumar, P., Kim, K.H., Lee, S.S., Bhattacharya, S.S., 2019. Solid waste management: scope and the challenge of sustainability. J. Clean. Prod. 228, 658–678.

Fell, D., Cox, J., Wilson, D.C., 2010. Future waste growth, modelling and decoupling. Waste Manag. Res. 28 (3), 281–286.

Fletcher, R., Rammelt, C., 2017. Decoupling: a key fantasy of the Post-2015 sustainable development agenda. Globalizations 14 (3), 450–467.

Gómez-Sanabria, A., Kiesewetter, G., Klimont, Z., Schoepp, W., Haberl, H., 2022. Potential for future reductions of global GHG and air pollutants from circular waste management systems. Nat. Commun. 13 (1).

Hapfelmeier, A., Ulm, K., 2013. A new variable selection approach using Random Forests. Comput. Stat. Data Anal. 60 (1), 50–69.

Hoornweg, D., Bhada-Tata, P., Kennedy, C., 2013. Waste production must peak this century. Nature 502 (7473), 615–617.

Jaligot, R., Chenal, J., 2018. Decoupling municipal solid waste generation and economic growth in the canton of Vaud, Switzerland. Resour. Conserv. Recycl. 130, 260–266.

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L., 2015. Plastic waste inputs from land into the ocean. Science 347 (6223), 768–771.

Kawai, K., Tasaki, T., 2016. Revisiting estimates of municipal solid waste generation per capita and their reliability. J. Mater. Cycl. Waste Manag. 18 (1), 1–13.

Kaza, S., Yao, L., Bhada-Tata, P., Woerden, F.Van, 2018a. What a Waste Global Database - City Level Dataset. The World Bank Group.

Kaza, S., Yao, L.C., Perinaz, B.-T., Woerden, F.Van, 2018b. What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050. The World Bank Group, Washington, D.C.

Kaza, S., Shrikanth, S., Chaudhary, S., 2021. More growth, less garbage. In: Bank, W. (Ed.), Urban Development Series. 47. World Bank, Washington D.C.

Kummu, M., Taka, M., Guillaume, J.H.A., 2018. Gridded global datasets for gross domestic product and human development index over 1990–2015. Scientific Data 5 (1), 180004.

Lau, W.W.Y., Shiran, Y., Bailey, R.M., Cook, E., Stuchtey, M.R., Koskella, J., Velis, C.A., Godfrey, L., Boucher, J., Murphy, M.B., Thompson, R.C., Jankowska, E., Castillo, A. Castillo, Pilditch, T.D., Dixon, B., Koerselman, L., Kosior, E., Favoino, E., Gutberlet, J., Baulch, S., Atreya, M.E., Fischer, D., He, K.K., Petit, M.M., Sumaila, U.R., Neil, E., Bernhofen, M.V., Lawrence, K., Palardy, J.E., 2020. Evaluating scenarios toward zero plastic pollution. Science (New York, N.Y.) 369 (6510), 1455–1461.

Liu, Y., Kong, F., Santibanez Gonzalez, E.D.R., 2017. Dumping, waste management and ecological security: evidence from England. J. Clean. Prod. 167, 1425–1437.

Macrotrends, 2020. European Union GDP Per Capita 1966-2020. 2020.

Madden, B., Florin, N., Mohr, S., Giurco, D., 2019. Using the waste Kuznet's curve to explore regional variation in the decoupling of waste generation and socioeconomic indicators. Resour. Conserv. Recycl. 149, 674–686.

Mazzanti, M., Zoboli, R., 2008. Waste generation, waste disposal and policy effectiveness. Evidence on decoupling from the European Union. Resour. Conserv. Recycl. 52 (10), 1221–1234.

OECD, 2020. OECD data: Municipal waste. OECD Data. 2020. OECD.org OECD.

Preston, F., Lehne, J., 2017. A Wider Circle? The Circular Economy in Developing Countries. Energy, Environment and Resources Department. The Royal Institute of International Affiars, London, Chattam House, p. 24.

Reyna-Bensusan, N., Wilson, D.C., Davy, P.M., Fuller, G.W., Fowler, G.D., Smith, S.R., 2019. Experimental measurements of black carbon emission factors to estimate the global impact of uncontrolled burning of waste. Atmos. Environ. 213, 629–639.

Solid waste management in the world's cities. In: Scheinberg, A., Wilson, D.C., Rodic-Wiersma, L. (Eds.), UN-Habitat;s State of Water and Sanitation in the World's Cities. Earthscan for UN-Habitat, London and Washington DC.

Strasser, S., 1999. Waste and Want: A Social History of Trash. Owl Books, New York.

Stretz, J., 2012. Economic Instruments in Solid Waste Management – Case Study Maputo, Mozambique. GIZ, Eshborn, Germany.

Stretz, J., 2013. Operator models: respecting diversity. Annex 7 - Maputo Case Study. GIZ, Eshborn, Germany.

Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9.

The Association of cities and regions for recycling and sustainable resource management (ACR+), 2009. Municipal Waste in Europe: Towards a European Recycling Society. ACR+.

Tisserant, A., Pauliuk, S., Merciai, S., Schmidt, J., Fry, J., Wood, R., Tukker, A., 2017. Solid waste and the circular economy: a global analysis of waste treatment and waste footprints. J. Ind. Ecol. 21 (3), 628–640.

UNEP, ISWA, 2015. Global waste management outlook. UNEP - International Environment Technology Centre. UNEP, ISWA, Osaka, Japan.

UN-Habitat, 2020. Waste Wise Cites TOOL - Step by Step Guide to Assess a City's Municipal Solid Waste Management Performance Through SDG Indicator 11.6.1 monitoring. 84. UN-Habitat - Supported by Wasteaware, EAWAG-SANDEC and University of Leeds, Nairobi.

UN-Habitat, n.d.UN-Habitat (n.d.). (Undated). "Waste Wise Cities Campaign." Retrieved 01 July 2020, from https://unhabitat.org/waste-wise-cities.

United Nations (UN), 2015. Transforming our world: the 2030 agenda for sustainable development. Sustainable Development Goals Knowledge Platform, UN.

United Nations (UN), 2018. World Urbanization Prospects: The 2018 Revision. online edition. United Nations Department of Economic and Social Affairs (DESA) Population Division.

Velis, C.A., 2017. Waste pickers in global south: informal recycling sector in a circular economy era. Waste Manag. Res. 35 (4), 329–331.

Weaver, R., 2015. Littering in context(s): using a quasi-natural experiment to explore geographic influences on antisocial behavior. Appl. Geogr. 57, 142–153.

Whiteman, A., Smith, P., Wilson, D.C., 2001. Waste management an indicator of urban governance. UN-Habitat: Global Habitat Conference on Urban Development. New York.

Whiteman, A., Webster, M., Wilson, D.C., 2021. The nine development bands: a conceptual framework and global theory for waste and development. Waste Manag. Res. 39 (10), 1218–1236.

Wiedinmyer, C., Yokelson, R.J., Gullett, B.K., 2014. Global emissions of trace gases, particulate matter, and hazardous air pollutants from open burning of domestic waste. Environ. Sci. Technol. 48 (16), 9523–9530.

Wilson, D.C., 2007. Development drivers for waste management. Waste Manag. Res. 25 (3), 198–207.

Wilson, D.C., Rodic, L., Scheinberg, A., Velis, C.A., Alabaster, G., 2012. Comparative analysis of solid waste management in 20 cities. Waste Manag. Res. 30 (3), 237–254.

Wilson, D.C., Rodic, L., Cowing, M.J., Velis, C.A., Whiteman, A.D., Scheinberg, A., Vilches, R., Masterson, D., Stretz, J., Oelz, B., 2015b. Supporting information to: 'Wasteaware' benchmark indicators for integrated sustainable waste management in cities. Waste Manag. 35.

Wilson, D.C., Rodic, L., Cowing, M.J., Velis, C.A., Whiteman, A.D., Scheinberg, A., Vilches, R., Masterson, D., Stretz, J., Oelz, B., 2015a. 'Wasteaware' benchmark indicators for integrated sustainable waste management in cities. Waste Manag. 35, 329–342.