FINAL MANUSCRIPT

# The efficient, the intensive, and the productive: Insights from urban Kaya scaling

Ramana Gudipudi[a], Diego Rybski[a,*], Matthias K.B. Lüdeke[a], Bin Zhou[a,b], Zhu Liu[c,d,e], Jürgen P. Kropp[a,f]

[a] Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany
[b] Desert Architecture and Urban Planning Unit, Swiss Institute for Dryland Environmental and Energy Research, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Israel
[c] Department of Earth System Science, Tsinghua University, Beijing 100084, China
[d] Resnick Sustainability Institute, California Institute of Technology, Pasadena, CA 91125, USA
[e] Tyndall Centre and School of International Development, University of East Anglia, Norwich NR4TJ, UK
[f] Dept. of Geo- and Environmental Sciences, University of Potsdam, 14476 Potsdam, Germany

A B S T R A C T

Urban areas play an unprecedented role in potentially mitigating climate change and supporting sustainable development. In light of the rapid urbanisation in many parts on the globe, it is crucial to understand the relationship between settlement size and $CO_2$ emission efficiency of cities. Recent literature on urban scaling properties of emissions as a function of population size has led to contradictory results and more importantly, lacked an in-depth investigation of the essential factors and causes explaining such scaling properties. Therefore, in analogy to the well-established Kaya Identity, we develop a relation combining the involved exponents. We demonstrate that application of this *Urban Kaya Relation* will enable a comprehensive understanding about the intrinsic factors determining emission efficiencies in large cities by applying it to a global dataset of 61 cities. Contrary to traditional urban scaling studies which use Ordinary Least Squares (OLS) regression, we show that the Reduced Major Axis (RMA) is necessary when complex relations among scaling exponents are to be investigated. RMA is given by the geometric mean of the two OLS slopes obtained by interchanging the dependent and independent variable. We discuss the potential of the Urban Kaya Relation in mainstreaming local actions for climate change mitigation.

## 1. Introduction

Harbouring more than 50% of the global population [1], contemporary cities generate 80% of the Gross Domestic Product (GDP) while consuming approximately 70% of the energy supply and releasing approximately three quarters of global $CO_2$ emissions [2]. Their unprecedented scale and complexity led to the development of a science of cities [3]. Drawing parallels between the allometric scaling in biological systems to that of cities, it has been studied how certain socioeconomic and environmental indicators in cities scale as a function of city size by means of the *urban scaling* approach [4]. Since a large fraction of the global population is expected to live in cities by end of this century [5], contemporary and future cities will play a pivotal role in global sustainability and climate change mitigation. Given this strong global urbanisation trend, one of the crucial questions that needs to be addressed is whether large cities are more (or less) emission efficient in

* Corresponding author.
*E-mail address:* ca-dr@rybski.de (D. Rybski).

comparison to smaller cities.

With an objective to quantify the performance of urban areas, the application of urban scaling has triggered copious research in the contemporary science of cities. Research on urban scaling can be broadly divided into two categories. The first category addresses the socio-economic performance of urban areas such as crime [6], social structures [7], and urban innovation [8]. The second category aims at quantifying the performance of urban infrastructure such as road networks [9] and modelling geographical networks [10]. Some researchers further developed the so-called 'scale adjusted metrics' to quantify the performance of cities [11,12]. Urban scaling relates a city indicator (e.g. total urban energy consumption) with city size (e.g. population). Assuming power-law correlations, the analysis depicts how these indicators scale with population size and whether large cities are more or less efficient. A sub-linear scaling (i.e. exponent $\beta < 1$) indicates that large cities consume less, e.g., energy given their size, while a linear scaling ($\beta \simeq 1$) depicts proportionality, and a super-linear scaling ($\beta > 1$) indicates that large cities consume more energy given their size.

It is stated that the per capita emissions of (many) urban areas is less than their respective country level per capita emissions mostly as a result of their relatively efficient urban form, building and transport infrastructure [13]. However, there is no general consensus on whether large cities are more energy and emission efficient in comparison to smaller cities. For instance, for total $CO_2$ emissions from cities in the USA, one study reported almost linear scaling [14], while another one reported super-linear scaling [15]. A similar study for European cities depicted super-linear scaling [11]. Studies on household electricity consumption in Germany and Spain revealed an almost linear scaling [4,16]. With respect to energy consumed and the subsequent emissions from urban transportation at a household level in the USA, Glaeser & Kahn [17] found a sub-linear scaling between population size and gasoline consumption; while another study depicted a super-linear scaling of emissions with population size [18]. A similar study done on British cities [19] found a linear scaling between transport emissions and population size while finding a super-linear relationship between emissions and the total street length. While the contradiction in these results can be majorly attributed to different city definitions and different indicators used to analyse, we identified two crucial research gaps that needs to be addressed in order to have an in-depth understanding about urban emission scaling. Firstly, most of these studies fail to explain the underlying systematic dynamics that govern these scaling properties. Secondly, most of these studies are limited to the urban systems within a given country and therefore fail to give a global overview of the emission (in-) efficiency of large cities. Therefore, in this paper we address these issues by developing a framework and apply it on a global dataset of 61 cities from 12 different countries in order to understand the intrinsic factors that determine scaling properties of urban emissions.

This is achieved by transferring the idea of the well-established Kaya Identity to urban $CO_2$ emissions leading to an *Urban Kaya Relation.* Then the scaling of $CO_2$ emissions with city population size can be attributed to the scaling of GDP with population, energy with GDP, and emissions with energy (further details in Section 2). To the best of our knowledge, such an attempt to obtain a further insight into the scaling of emissions with population using indicators in the Kaya Identity is unprecedented. Recent literature has identified that the energy consumption and the subsequent emissions depend on the city type (i.e. affluent and mature cities in developed countries versus cities in transition countries with emerging and nascent infrastructure) [20,2]. Therefore, we apply the Urban Kaya Relation to these cities separately.

## 2. Urban Kaya relation

The *Kaya Identity* [21] can be written as

$$\frac{C}{P} = \frac{G}{P}\frac{E}{G}\frac{C}{E}, \tag{1}$$

where $C$ is the total emissions, $P$ is the population, $G$ is the GDP and, $E$ is the energy consumption. It is an identity since all the numerators and denominators cancel down to $C/P = C/P$. While the GDP per capita ($G/P$) is a common quantity for estimating affluence, the energy intensity ($E/G$) can be understood as the energy necessary to generate GDP, and the carbon intensity ($C/E$) as the efficiency in energy production and consumption (technological).

The urban scaling hypothesis states that certain city properties such as GDP and emissions exhibit scaling relationships as a function of population size. Therefore, we propose that each of the variables in numerator and denominator in Eq. (1) exhibit scaling, i.e.

$$C \sim P^{\phi} \tag{2}$$

$$G \sim P^{\beta} \tag{3}$$

$$E \sim G^{\alpha} \tag{4}$$

$$C \sim E^{\gamma}. \tag{5}$$

As outlined above, here we are interested in how the urban $CO_2$ emissions scale with population size, Eq. (2). The value of $\phi$ tells us if large or small cities are more efficient in terms of $CO_2$ emissions. Superlinearity of Eq. (3) with $\beta > 1$ is well known in agglomeration economics, see e.g. [22], and has recently been confirmed [4]. Equation (4) has been studied on the country scale [23]. The established power-law relations Eqs. (2)–(4) indicate that also Eq. (5) holds. In case the power-law form is not empirically supported, Eqs. (2)–(5) can still be considered as linear approximations (in log-log space) of potentially more complex functional forms.

Combining the scaling relations in Eq. (2)–(5) – e.g. $C \sim E^{\gamma}$ and $E \sim G^{\alpha}$ leads to $C \sim (G^{\alpha})^{\gamma} \sim G^{\alpha\gamma}$ – we get

$$\phi = \beta\,\alpha\,\gamma. \tag{6}$$

Thus, in analogy to the original Kaya Identity, we obtain the *Urban Kaya Relation*, Eq. (6) which structurally resembles to the traditional Kaya Identity in Eq. (1). According to Eq. (6), the exponent relating emissions and population is simply given by the product of the other involved exponents. This permits us to attribute non-linear scaling of emissions with city size (Eq. (2)), to potential urban scaling of GDP with population, energy with GDP, or emissions with energy. For the sake of completeness, in Appendix A we also provide another two complementary forms of Kaya Identities and corresponding Urban Kaya Relations.

However, the exponent $\phi$ is usually obtained from data and a linear regression $\ln C = \phi \ln P + a$, where $a$ is another fitting parameter. Equations (2)–(5) represent idealisations and in practice correlations are studied which can come with more or less spread around the regression. Ordinary Least Squares (OLS) might make sense, when dependent and independent variables are clearly defined, e.g. in the case of GDP vs. population it might be preferable to minimise residuals of GDP. Applying OLS to $C \sim P^{\phi}$ and $P \sim C^{1/\phi^*}$ generally leads to $\phi \neq \phi^*$ [24,25] so that also Eq. (6) would not hold (see Appendix B.1). In our context, however, dependent and independent variables need to be exchangeable and we obtained robust results ($\phi = \phi^*$) by applying *Reduced Major Axis* regression (RMA, see Appendix B.2). Therefore, we apply RMA throughout the paper unless specified otherwise. In RMA, the slope is given by the geometric mean of the two OLS slopes obtained by interchanging the dependent and independent variable [26,27]. The authors would like to emphasize that the usage of RMA in this study is only to ensure that Eq. (6) is formally valid. In order to quantify the uncertainty of the estimated exponents, we explore bootstrapping, applying 20,000 replications.

## 3. Data

The major pre-requisites while investigating the scaling effects of urban energy consumption and emissions are (a) a consistent definition and demarcation of cities from their hinterlands and (b) a consistent accounting approach to quantify the energy consumption and subsequent emissions [2]. The analysis conducted in this paper is limited to 61 global cities, i.e. the intersection of cities for which the 4 quantities are available, i.e. (i) $CO_2$ emissions, (ii) total final energy consumption, (iii) GDP, and (iv) population. Although, the data used in this analysis might be inconsistent owing to the challenges mentioned above, we used it as a showcase to demonstrate the applicability of the Urban Kaya Relation. The limitations of the data and its implications on the exploratory results are discussed in the Section 5.

The population, GDP, and total final energy consumption data used in this study is taken from the Chapter 18 "Urban Energy Systems" of the Global Energy Assessment [28]. This database includes the per capita total final energy consumption of 223 global cities, their respective population and GDP for the year 2005. The data on emissions is compiled from various sources including city specific reports (provided by organisations such as ICLEI [29], CDP [30], and C40 cities [31]) and data which is published in peer reviewed journals [32].

The cities with available data are located in 12 countries. The GDP per capita of these countries shows two groups. One ranging from 740 USD to 4700 USD and the other from 26000 USD to 44000 USD (year 2005). These two groups can be considered developing and developed countries and represent the Non-Annex 1 and Annex 1 countries as reported by the United Nations Framework Convention on Climate Change (UNFCCC), respectively. Amongst the 61 cities used in this analysis 22 cities are from the Annex I countries and 39 cities from Non-Annex I countries. The database consists of cities of varying population sizes across 6 continents including 7 mega-cities (with a reported population above 10 million). Within countries in Annex 1 regions, 7 cities in the USA, 4 cities in the UK, 2 cities in Germany, Spain, Australia, Italy, France, respectively, and 1 city in Japan are considered in this study. With respect to cities in Non-Annex 1 countries 33 cities in China, 2 cities in India, South Africa, and Brazil, respectively, were included.

On the country scale, $CO_2$ emissions per capita strongly depend on the development of the considered country, see e.g. [33] and references therein. Here we pool together cities from many different countries, including from developing countries; as a consequence, the data needs to be normalised prior to the analysis in order to account for baseline emissions, data and other inhomogeneities. Therefore, we employ a method that was recently proposed for urban scaling [11] and normalise the data for each country by the average logarithmic city size ($\langle \ln P \rangle$) and indicator value (e.g. $\langle \ln C \rangle$), whereas for each indicator we take the maximum available sample size.

## 4. Results

We begin by looking at the scaling of emissions with population size for the considered 61 cities. The slope of this logarithmic RMA (see Fig. 1) is almost equal to one ($\phi = 1.01$), however, the pattern of residuals is diverse as also reported in some earlier studies [11]. This result shows that at a global scale large cities are typically not more emission efficient compared to smaller cities. Further, in Fig. 1 a distinction between the cities in developed countries (Annex 1) and cities in developing countries (Non-Annex 1) is made.

For comparison, in Table 1 we list the resulting exponents, when we employ OLS to the scaling of the 61 cities. Table 1 also includes the absolute difference between the prediction (Eq. (6)) and the measured exponent $\phi$. The obtained exponents deviate strongly when OLS is used (instead of RMA). As discussed at the end of Section 2, we attribute this discrepancy to different regressions when minimising the residuals along any of both axes. In the case of RMA, plotting $G$ vs. $P$ and $P$ vs. $G$

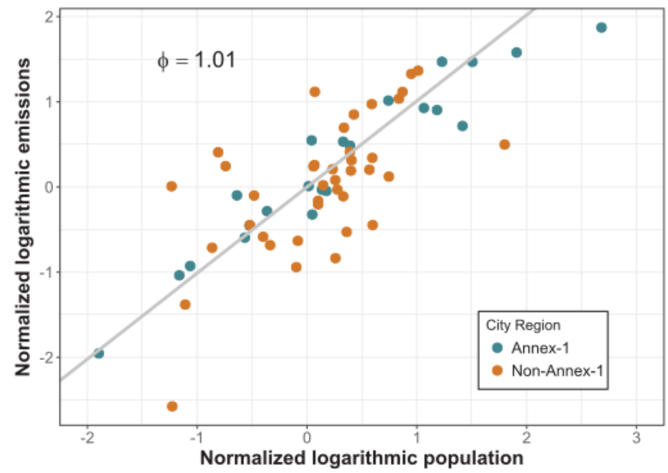


**Fig. 1.** Scaling of $CO_2$ emissions with population size for 61 global cities. The data has been normalised by subtracting average logarithmic values, see Section 3. The solid grey line is the regression obtained from the *Reduced Major Axis* (RMA, Section 2) and the slope is $\phi \simeq 1.01$, see Table 1 for details.

leads by definition to the same result. Thus, we recommend to employ RMA instead of OLS when studying the Urban Kaya Relation. Moreover, for OLS it has been shown that whether the estimated exponents are statistically different from 1 depends on the assumptions made [34].

As a next step, we analysed the scaling properties of emissions with size separately depending on the economic geography of the country (i.e. Annex 1 cities vs. Non-Annex 1) in which these cities are located. In Fig. 2 we see that the scaling of emissions with the population size indeed has a dependence on the economic geography of the country. We found a sub-linear scaling for cities in Annex 1 regions ($\phi = 0.87$) and a super-linear scaling for cities in the Non-Annex 1 regions ($\phi = 1.18$), see Table 1. In order to test if these slopes are significantly different, we perform bootstrapping and a Kolmogorov-Smirnov (KS) test. The KS distance between these bootstrapped samples is 0.83 with a significant P-value ($< 2.2 \times 10^{-16}$) which confirms that the slopes are not drawn from the same distributions. The fit appears to be good for cities in Annex 1 regions which are broadly characterised as service sector oriented economies. However, in industry dominated Non-Annex 1 cities with widely varying infrastructure and energy intensity of production the goodness of fit appears to be relatively poor. This result shows either that the emissions data from Non-Annex 1 cities is not as accurate, or that population is a good proxy to estimate emissions for cities in Annex 1 regions while there seems to be other factors that influence emissions for cities in Non-Annex 1 countries.

We looked at scaling of each of the indicators in the Urban Kaya Relation, namely the scaling of GDP with population ($G/P$) Eq. (3), scaling of energy intensity ($E/G$) Eq. (4), and carbon intensity ($C/E$) Eq. (5). Table 1 lists the exponents for each of these relations. From a global perspective, our results suggest that the almost linear scaling of emissions with population size could be attributed to the almost linear scaling of carbon intensity and the trade-off between scaling of GDP with population and the scaling of energy intensity (i.e. they compensate each other).

In the case of cities in Annex 1 countries, our results show that the large cities typically have lower emissions per capita compared to smaller cities because of the sub-linear scaling of the carbon intensity (Table 1). This might be attributed to the carbon intensity of the electricity generation supply mix, vehicle fuel economy, and the quality of public transit in these cities [35]. We found an approximately linear scaling of GDP with population. Our result shows that doubling the GDP in these cities will lead to an almost similar increase in energy consumption. Such a linear scaling might be largely attributed to the consumption patterns and infrastructure lock-in behaviour in largely

**Table 1**
Scaling exponents and Urban Kaya Relation. The Table lists the various estimated exponents and the last column shows how well the Urban Kaya Relation performs. The exponents are listed for all cities, cities in Annex 1 countries, and cities in Non-Annex 1 countries (see Section 3). All exponents have been obtained from RMA (see Section 2) except for the last row, where OLS has been applied for comparison. The square brackets give 95% confidence intervals from bootstrapping (20,000 replications). Inspired by the notation used in [34] we put the following symbols. ↗, at least 66.6% of the replications lead to exponents larger than 1;→, 33.3% to 66.6% of the estimates are larger than 1, and ↘, less than 33.3% are larger than 1. While Eq. (6) works exactly for RMA (see Appendix C2), for OLS the estimated exponents are incompatible (last row).

| Exponent: | $\phi$ | $\beta$ | $\alpha$ | $\gamma$ | $\lvert\phi - \beta\,\alpha\,\gamma\rvert$ |
|---|---|---|---|---|---|
| Equation: | Eq. (2) | Eq. (3) | Eq. (4) | Eq. (5) | Eq. (6) |
| Scaling of: | Emissions with population | GDP with population | Energy with GDP | Emissions with Energy | |
| All Cities | 1.01 → [0.87,1.18] | 1.13 ↗ [1.04,1.24] | 0.92 ↘ [0.78,1.06] | 0.97 → [0.82,1.14] | 0.00 |
| Annex 1 | 0.87 ↘ [0.59,0.95] | 1.03 ↗ [1.00,1.15] | 0.99 → [0.60,1.63] | 0.85 ↘ [0.41,1.21] | 0.00 |
| Non-Annex 1 | 1.18 ↗ [0.90,1.49] | 1.27 ↗ [1.04,1.55] | 0.83 ↘ [0.67,1.00] | 1.11 ↗ [0.93,1.29] | 0.00 |
| All Cities (OLS) | 0.80 | 0.98 | 0.64 | 0.67 | 0.38 |

service based economies [36,20].

We further checked if the sub-linear scaling of emissions with population for cities in Annex 1 countries could be attributed to a possible sub-linear scaling with respect to their total final energy consumption Eq. (A.3). Even in a completely decarbonised world, the question of energy efficiency will persist. Our results suggest that large cities in Annex 1 countries are not much more energy efficient with respect to their population ($\delta \simeq 1.04$, see Appendix A) compared to smaller cities. This result indicates that although the per capita energy consumption in large cities is similar to that of smaller cities, it is the better technologies employed in larger cities that typically make their per capita emissions lower than those of smaller cities.

With respect to cities in Non-Annex 1 countries, our results suggest that the super-linear scaling of emissions with population is due to two factors: (1) super-linear scaling of GDP with population and (2) super-linear scaling of carbon intensity. However, we found that doubling the GDP in these cities will lead to a less than double increase in energy consumption. This might be attributed to the prevalence of energy poverty in these cities [36]. Large cities in Non-Annex 1 countries benefit economically (more GDP) from the urban poor who consume less energy and have limited access to electricity. Therefore, large cities in this region are more energy efficient compared to smaller ones.

## 5. Discussion

The impact of urbanization on energy consumption and subsequent emissions is of particular interest considering climate change and mitigation ambitions. Previous studies attempted to quantify this impact

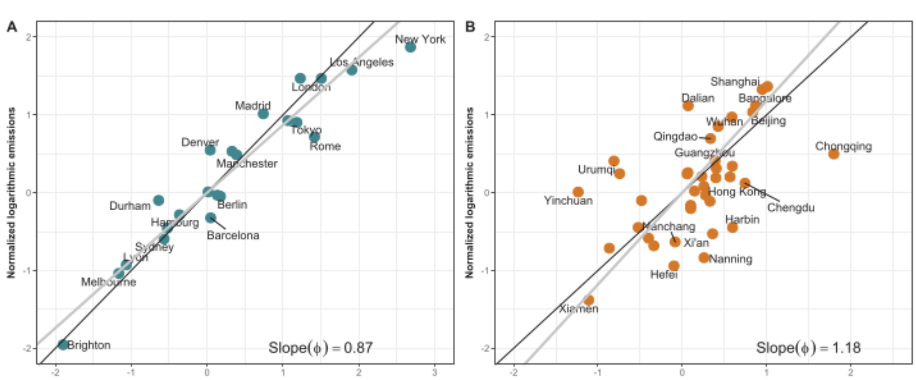


**Fig. 2.** Scaling of $CO_2$ emissions with population for cities in Annex 1 countries (panel A) and in Non-Annex 1 countries (panel B). While the slope of the RMA (grey) for Annex 1 countries is found to be sub-linear ($\phi \simeq 0.87$), it is super-linear ($\phi \simeq 1.18$) with respect to cities in Non-Annex 1 countries. As in Fig. 1 the data of both axes has been normalised subtracting average logarithmic values, see Section 3. The black line indicates a slope of 1 and is included for comparison.

at country scale using the IPAT equation (see Appendix C). For instance, one study concluded that urbanization decreases energy consumption in low-income countries while increases emissions in middle income groups [37]. Another cross-country panel data analysis revealed that the impact of urbanization on emissions follows a pattern similar to the environmental Kuznets curve with a tipping point at 73.80% urbanization rate [38]. While these studies give an overview of the impact of urbanization on emissions at a country scale, there is a need to address two issues which are crucial to streamline mitigation strategies at city scale. Firstly, these studies do not state whether large cities are more (in) efficient in comparison to smaller cities. Secondly, unlike 'Kaya Identity', the choice of independent variable for 'technology' varied in studies using the IPAT equation to quantify emissions. We address these two issues in this paper by combining two well-established concepts namely 'Urban Scaling' with 'Kaya Identity' and developed the 'Urban Kaya Relation'. We believe that such a formal foundation is crucial to identify the factors contributing to emission (in-) efficiency of large cities and streamline local actions for climate change mitigation. Moreover, the exponents in the Urban Kaya Relation ($\phi$, $\beta$, $\alpha$ and $\gamma$) are fundamentally different from those within stochastic IPAT and traditional Kaya Identity as these exponents are the scaling exponents of the variables in respective numerator and denominators (emissions/population, GDP/population, energy/GDP and emissions/energy respectively).

The urban scaling approach can be attributed to city functionality as a (short-term) spatial equilibrium of the interplay between the density dependent socioeconomic interactions and transportation costs [11]. The universality of the scaling properties of socioeconomic (superlinear) and infrastructural components (sub-linear) in cities can be explained as an outcome of the increased interaction between citizens at a microscopic scale [39] and it's combination with the fractal properties of cities [40]. However, unlike socioeconomic and infrastructure components, urban emissions are a result of various economic activities (e.g. location of industries). Such processes consume resources and energy which are often originated beyond the urban boundaries but have an implication on emissions from electricity consumption. In such cases, population alone might not be a good predictor for emissions. This could be the explanation for the relatively poor fitting of the emission scaling properties of cities in Non-Annex 1 regions in comparison to cities in Annex 1 regions.

The data used in this study has three major constraints. Firstly, the sources for emission data is different from the source of energy consumption and GDP. Therefore, the urban extents and the population size might vary for few cities. Studies have shown that such city definitions will influence the scaling properties [41,42]. Secondly, since the data is from multiple sources, the accounting approach varied in most of the cities. Inclusion (or exclusion) of emissions embedded in electricity generation will have a significant impact on energy and emissions attributed to building sector. Thirdly, the sectoral emissions and type of fuels used varies from one city to another. The energy data from the GEA [28] study includes the total final energy consumption but excludes traditional biomass for few cities in developing regions. Data on emissions attributed to the industrial sector is inconsistent as it excludes industrial electricity and fuel usage in some cities. In the view of aforementioned data limitations, the broader conclusions drawn in this paper should be treated cautiously since the empirical part is rather intended as a proof of concept.

Further, we need to mention the role of urban population *density* as an important factor in determining the energy consumption and subsequent emissions. On the one hand, it has been shown that urban $CO_2$ emissions from transport energy per capita decrease with population density [35,20,43]. On the other hand, there is a theoretical connection between urban indicators, population, and area scalings [44,25]. Combining density with the Urban Kaya Relation introduces further complexity which we leave to be addressed by future research.

## 6. Conclusions

In summary, the achievements of this work are threefold. (i) In analogy to the Kaya Identity – which in the climate change community represents a well know specification of the IPAT approach (see Appendix C) – we set out a framework to assess why urban $CO_2$ emissions scale super- or sub-linearly with city size (Eq. (2)). We derive the *Urban Kaya Relation* $\phi = \beta\alpha\gamma$. (ii) We show that Ordinary Least Squares (OLS) lead to inconsistent results and propose to use *Reduced Major Axis* (RMA) regression. (iii) As a proof of concept we apply the Kaya framework to the available data. In the first place, the proposed Kaya relation can be used to see from which (in)efficiency $\phi \neq 1$ is stemming from. In the second place, it can serve as consistency check, i.e. the product of exponents must be correct.

It is crucial to establish foundations in the form of such a framework to understand the guiding factors that govern scaling properties since urban areas are often identified as the focal spatial units for improving energy efficiency and climate change mitigation [13,45]. Urban energy consumption and subsequent emissions are an outcome of urbanites' affluence and their consumption patterns [46]. Nevertheless, it is important to investigate whether the infrastructural efficiency of large cities will be manifested as emission efficiency gains. An in-depth investigation about the demographic, economic and technological drivers of urban emissions is necessary to identify the key entry points for mitigation actions at a city scale. By means of an exploratory analysis we demonstrate that the Urban Kaya Relation can be used to address this issue adequately by attributing the scaling properties of emissions to the scaling of GDP with population (affluence), energy intensity (economic geography) and emission intensity (technology).

Our exploratory results show that large cities in Annex 1 countries have lower emissions compared to smaller cities. This result suggests the usage of better technologies in energy generation/consumption and efficient modes of transportation. From a climate change mitigation point of view, the key challenge in these cities is to further decrease their energy and carbon intensity while ensuring economic stability. According to our exploratory results larger cities in emerging countries such as China, India, and Brazil typically have more per capita emissions compared to smaller cities. From one point of view, it may be good news that large cities in these regions are not emission efficient since much of the urbanisation in these regions is going to happen in small and medium size cities [47]. Thus, despite being exploratory, our findings corroborate the results of previous studies which showed the significance of affluence on emissions [48] and the influence of economic geography on the scaling properties of emissions with population [25]. Further support comes from a recent study, where a methodology other than urban scaling has been applied and completely different data has been used [49].

The data constraints mentioned in the previous chapter also highlight the data needs. Therefore, we acknowledge the ongoing efforts to develop a consistent emission framework by various international organisations[1] and make an appeal that such efforts should disclose data on energy consumption and GDP along with sectoral emissions and population. It is envisaged that the application of Urban Kaya Relation to a consistent dataset will significantly contribute to the discussion whether large cities are more (or less) emission efficient and the underlying scaling relations. Since cities are already taking measures to improve energy and emission efficiency as focal points for climate change mitigation, our results suggest that cities in Annex 1 regions should primarily focus on improving their energy efficiency while cities in the Non-Annex 1 regions should focus on a technological shift towards providing universal access to energy and deploy more energy efficient technologies.

[1] http://www.ghgprotocol.org/.

## Data accessibility

All data used are publicly available, sources are referenced in Section 3.

## Authors' contributions

RG, DR, MKBL, BZ, JPK designed the study; ZL, RG collected the data; RG carried out the data analysis; RG, DR, MKBL, BZ validated the results; RG prepared the Figures; RG, DR, MKBL, JPK wrote and revised the manuscript.

## Competing interests

We declare we have no competing interests.

## Research ethics

No experiments were carried out during this study and no prior ethical assessment was required.


## Funding

The research leading to these results has received funding from the European Community's Seventh Framework Programme under Grant Agreement No. 308497 (Project RAMSES). Zhu Liu acknowledges support by NSFC (71874079, 41501605) and the Green Talents Program held by the German Federal Ministry of Education and Research (BMBF).

## Acknowledgment

We thank M. Barthelemy, H.V. Ribeiro, and L. Costa for useful discussions. This work emerged from ideas discussed at the symposium *Cities as Complex Systems* (Hanover, July 13th-15th, 2016) which was generously funded by Volkswagen Foundation.


## Appendix A. Kaya II and III

It needs to be mentioned that there are another two identities complementary to the original Kaya Identity, Eq. (1), namely

$$C = P\frac{E}{P}\frac{G}{E}\frac{C}{G} \tag{A.1}$$

$$C = G\frac{P}{G}\frac{E}{P}\frac{C}{E}, \tag{A.2}$$

or variations. We propose to denote Eqs. (1), (A.1), and (A.2), "Kaya I", "Kaya II", and "Kaya III', respectively. The identities Kaya II and III involve two intensities which do not appear in Kaya I, namely $E/P$ and $C/G$, i.e. energy per capita and carbon per GDP, respectively. In the urban scaling picture these take the form

$$E \sim P^{\delta} \tag{A.3}$$

$$C \sim G^{\eta}.$$

The relations corresponding to Eqs. (A.1) and (A.2) are

$$\phi = \frac{\delta\eta}{\alpha}$$

$$\eta = \frac{\delta\gamma}{\beta}.$$

Other combinations of $C$, $P$, $G$, or $E$ involve only two components each.

## Appendix B. Different linear regression slopes

### B.1. Ordinary Least Squares (OLS)

If we consider $c_i = \ln C_i$ and $p_i = \ln P_i$ with standard deviations $\sigma_c$ and $\sigma_p$, respectively, then the slope according to $\mathrm{OLS}(c|p)$ is analytically given by

$$\phi = \rho_{c,p}\frac{\sigma_c}{\sigma_p} \tag{B.1}$$

where $\rho_{c,p}$ is the correlation coefficient [26,27]. Accordingly, for Eq. (6) we obtain

$$\beta\alpha\gamma = \rho_{g,p}\frac{\sigma_g}{\sigma_p}\rho_{e,g}\frac{\sigma_e}{\sigma_g}\rho_{c,e}\frac{\sigma_c}{\sigma_e} \tag{B.2}$$

$$= \rho_{g,p}\rho_{e,g}\rho_{c,e}\frac{\sigma_c}{\sigma_p}. \tag{B.3}$$

Comparison with Eq. (B.1) leads to $\rho_{c,p} = \rho_{g,p}\rho_{e,g}\rho_{c,e}$, which is *not true* in general.

### B.2. Reduced Major Axis (RMA)

The Reduced Major Axis (RMA) is given by the geometric mean of the two OLS slopes, i.e. minimizing the sum of squares of vertical residuals, $\mathrm{OLS}(y|x)$, or horizontal residuals, $\mathrm{OLS}(x|y)$, respectively [26,27]. Then the slope according to RMA is analytically given by

$$\phi = \mathrm{sign}(\rho_{c,p})\frac{\sigma_c}{\sigma_p}. \tag{B.4}$$

Since in our case the correlations are always positive, we omit $\mathrm{sign}(\rho)$ from now on. Then, for Eq. (6) we obtain

$$\beta\,\alpha\,\gamma = \frac{\sigma_g}{\sigma_p}\,\frac{\sigma_e}{\sigma_g}\,\frac{\sigma_c}{\sigma_e} \tag{B.5}$$

$$= \frac{\sigma_c}{\sigma_p}, \tag{B.6}$$

which is consistent with Eq. (B.4). One can see that the result is independent from any correlation coefficient.

Please note, in a previous version of our manuscript we used Orthogonal Regression (also known as Total Least Squares, TLS). Since the corresponding analytical expression for the slope is much more complex, here we employ the simpler RMA.

## Appendix C. IPAT concept

The original Kaya Identity is a specific version of the IPAT concept, which stands for

$$I = P\,A\,T, \tag{C.1}$$

where the quantities are impact ($I$), population ($P$), affluence ($A$), technology ($T$), see e.g. [50] and references therein. A stochastic IPAT version introduced in [51] is in this context given by

$$C = a\,P^b\,G^c\,E^d, \tag{C.2}$$

where $a\ldots d$ are parameters.

While Eq. (C.2) is a higher-dimensional extension of Eq. (2), with the goal of better predicting $C = \mathrm{f}(P,\,G,\,E)$, our approach is still based on Eq. (2) but aims at circumscribing it employing the other scaling relations Eqs. (3)–(5).

## References


[1] United Nations, World Urbanization Prospects 2014, Tech. rep.; 2014. https://doi.org/ST/ESA/SER.A/366.
[2] Seto KC, Dhakal S, Bigio A, Blanco H, Delgado GC, Dewar D, et al. Human settlements, infrastructure and spatial planning. Tech. rep. IPCC; 2014.
[3] Batty M. Building a science of cities. Cities 2012;29:S9–16. https://doi.org/10.1016/j.cities.2011.11.008.
[4] Bettencourt LMA, Lobo J, Helbing D, Kühnert C, West GB. Growth, innovation, scaling, and the pace of life in cities. Proc Nat Acad Sci USA 2007;104(17):7301–6. https://doi.org/10.1073/pnas.0610172104.
[5] Batty M. Commentary – when all the world's a city. Environ Plan A 2011;43(4):765–72. https://doi.org/10.1068/a43403.
[6] Alves LGA, Ribeiro HV, Lenzi EK, Mendes RS. Distance to the scaling law: a useful approach for unveiling relationships between crime and urban metrics. PLoS One 2013;8(8):e69580. https://doi.org/10.1371/journal.pone.0069580.
[7] Pan W, Ghoshal G, Krumme C, Cebrian M, Pentland A. Urban characteristics attributable to density-driven tie formation. Nat Com 2013;4:1961. https://doi.org/10.1038/ncomms2961.
[8] Arbesman S, Kleinberg JM, Strogatz SH. Superlinear scaling for innovation in cities. Phys Rev E 2009;79(1):016115. https://doi.org/10.1103/PhysRevE.79.016115.
[9] Samaniego H, Moses ME. Cities as organisms: allometric scaling of urban road networks. J Transp Land Use 2008;1(1):21–39. https://doi.org/10.5198/jtlu.v1i1.29.
[10] Yakubo K, Saijo Y, Korošak D. Superlinear and sublinear urban scaling in geographical networks modeling cities. Phys Rev E 2014;90(2):022803. https://doi.org/10.1103/PhysRevE.90.022803.
[11] Bettencourt LMA, Lobo J. Urban scaling in Europe. J R Soc Interface 2016;13(116):20160005. https://doi.org/10.1098/rsif.2016.0005.
[12] Alves LGA, Mendes RS, Lenzi EK, Ribeiro HV. Scale-adjusted metrics for predicting the evolution of urban indicators and quantifying the performance of cities. PLoS One 2015;10(9):e0134862. https://doi.org/10.1371/journal.pone.0134862.
[13] Dodman D. Blaming cities for climate change? An analysis of urban greenhouse gas emissions inventories. Environ. Urban. 2009;12(1):185–201. https://doi.org/10.1177/0956247809103016.
[14] Fragkias M, Lobo J, Strumsky D, Seto KC. Does size matter? Scaling of CO2 emissions and U.S. urban areas. PLoS One 2013;8(6):e64727. https://doi.org/10.1371/journal.pone.0064727.
[15] Oliveira EA, Andrade JS, Makse HA. Large cities are less green. Sci Rep 2014;4(4235):srep04235. https://doi.org/10.1038/srep04235.
[16] Horta-Bernús R, Rosas-Casals M. Obsolescence in urban energy infrastructures: the influence of scaling laws on consumption forecasting. J Urban Technol 2015;22(2):3–17. https://doi.org/10.1080/10630732.2014.955340.
[17] Glaeser EL, Kahn ME. The greenness of cities: carbon dioxide emissions and urban development. J Urban Econ 2010;67(3):404–18. https://doi.org/10.1016/j.jue.2009.11.006.
[18] Louf R, Barthelemy M. How congestion shapes cities: from mobility patterns to scaling. Sci Rep 2014;4(5561):srep05561. https://doi.org/10.1038/srep05561.
[19] Mohajeri N, Gudmundsson A, French JR. CO2 emissions in relation to street-network configuration and city size. Transp Res Part D Transp Environ 2015;35:116–29. https://doi.org/10.1016/j.trd.2014.11.025.
[20] Creutzig F, Baiocchi G, Bierkandt R, Pichler P-P, Seto KC. Global typology of urban energy use and potentials for an urbanization mitigation wedge. Proc Nat Acad Sci USA 2015;112(20):6283–8. https://doi.org/10.1073/pnas.1315545112.
[21] Raupach MR, Marland G, Ciais P, Quéré CL, Canadell JG, Klepper G, et al. Global and regional drivers of accelerating CO2 emissions. Proc Nat Acad Sci USA 2007;104(24):10288–93. https://doi.org/10.1073/pnas.0700609104.
[22] Sveikauskas L. The productivity of cities. Q J Econ 1975;89(3):393–413 <http://www.jstor.org/stable/1885259>.
[23] Brown JH, Burnside WR, Davidson AD, DeLong JP, Dunn WC, Hamilton MJ, et al. Energetic limits to economic growth. Bioscience 2011;61(1):19–26. https://doi.org/10.1525/bio.2011.61.1.7.
[24] Fluschnik T, Kriewald S, Ros AGC, Zhou B, Reusser DE, Kropp JP, Rybski D. The size distribution, scaling properties and spatial organization of urban clusters: a global and regional percolation perspective. Int J Geo-Inform 2016;5(7):110. https://doi.org/10.3390/ijgi5070110.
[25] Rybski D, Reusser DE, Winz AL, Fichtner C, Sterzel T, Kropp JP. Cities as nuclei of sustainability? Environ Plan B 2017;44(3):425–40. https://doi.org/10.1177/0265813516638340.
[26] Isobe T, Feigelson ED, Akritas MG, Babu GJ. Linear regression in astronomy. I. Astrophys J 1990;364(20):104–13. https://doi.org/10.1086/169390.
[27] Babu GJ, Feigelson ED. Analytical and monte carlo comparisons of six different linear least squares fits. Commun Statist -Simula 1992;21(2):533–49. https://doi.org/10.1080/03610919208813034.
[28] Grubler A, Bai X, Buettner T, Dhakal S, Fisk DJ, Ichinose T, et al. Urban energy systems. In: Global Energy Assessment - Toward a Sustainable Future, Cambridge University Press, Cambridge, UK and New York, NY, USA and the International Institute for Applied Systems Analysis, Laxenburg, Austria; 2013. p. 1307–400 [chapter 18].
[29] ICLEI, Energy and carbon emissions profiles of 54 south asian cities, Tech. rep., ICLEI - Local Governments for Sustainability, New Delhi (2009). URL <http://southasia-old.iclei.org/fileadmin/user_upload/documents/Energy_and_Carbon_Emissions_Profiles_for_54_South_Asian_Cities.pdf>.
[30] CDP Worldwide, CDP open data portal (2016). URL <https://data.cdp.net/>.
[31] C40 Cities, C40 Greenhouse Gas Protocol for Cities Interactive Dashboard; 2014. URL < http://www.c40.org/research/open_data/5 >.
[32] Shan Y, Guan D, Liu J, Liu Z, Liu J, Schroeder H, et al. CO2 emissions inventory of Chinese cities. Atmos Chem Phys Discuss 2016;2016:1–26. https://doi.org/10.5194/acp-2016-176.
[33] Costa L, Rybski D, Kropp JP. A human development framework for CO2 reductions. PLoS One 2011;6(12):e29262. https://doi.org/10.1371/journal.pone.0029262.
[34] Leitão JC, Miotto JM, Gerlach M, Altmann EG. Is this scaling nonlinear? R Soc Open Sci 3(7). https://doi.org/10.1098/rsos.150649.
[35] Kennedy C, Steinberger J, Gasson B, Hansen Y, Hillman T, Havránek M, et al. Greenhouse gas emissions from global cities. Environ Sci Technol 2009;43(19):7297–302. https://doi.org/10.1021/es900213p.
[36] Satterthwaite D. The implications of population growth and urbanization for climate change. Environ Urban 2009;21(2):545–67. https://doi.org/10.1177/0956247809344361.
[37] Poumanyvong P, Kaneko S. Does urbanization lead to less energy use and lower CO2 emissions? A cross-country analysis. Ecol Econ 2010;70(2):434–44. https://doi.org/10.1016/j.ecolecon.2010.09.029.
[38] Zhang N, Yu K, Chen Z. How does urbanization affect carbon dioxide emissions? A cross-country panel data analysis. Energy Policy 2017;107:678–87. https://doi.org/10.1016/j.enpol.2017.03.072 <http://linkinghub.elsevier.com/retrieve/pii/S0301421517301933>.
[39] Schläpfer M, Bettencourt LMA, Grauwin S, Raschke M, Claxton R, Smoreda Z, et al. The scaling of human interactions with city size. J R Soc Interface 2014;11(98):20130789. https://doi.org/10.1098/rsif.2013.0789.
[40] Ribeiro FL, Meirelles J, Ferreira FF, Neto CR. A model of urban scaling laws based on distance dependent interactions. R Soc Open Sci 2017;4(3):160926. https://doi.

org/10.1098/rsos.160926.

[41] Arcaute E, Hatna E, Ferguson P, Youn H, Johansson A, Batty M, et al. Constructing cities, deconstructing scaling laws. J R Soc Interface 2014;12(102):20140745. https://doi.org/10.1098/rsif.2014.0745.

[42] Cottineau C, Hatna E, Arcaute E, Batty M. Diverse cities or the systematic paradox of urban scaling laws. Comput Environ Urban Syst 2017;63:80–94. https://doi.org/10.1016/j.compenvurbsys.2016.04.006.

[43] Gudipudi R, Fluschnik T, Ros AGC, Walther C, Kropp JP. City density and CO2 efficiency. Energy Policy 2016;91:352–61. https://doi.org/10.1016/j.enpol.2016.01.015.

[44] Rybski D. Relating urban scaling, fundamental allometry, and density scaling, arXiv.org e-Print archive arXiv:1609.01217 [physics.soc-ph]. URL <https://arxiv.org/abs/1609.01217>.

[45] Hoornweg D, Sugar L, Gomez CLT. Cities and greenhouse gas emissions: moving forward. Environ Urban 2011;23(1):207–27. https://doi.org/10.1177/0956247810392270.

[46] Satterthwaite D. Cities' contribution to global warming: notes on the allocation of greenhouse gas emissions. Environ Urban 2008;20(2):539–49. https://doi.org/10.1177/0956247808096127.

[47] Department of Economic and Social Affairs, World Urbanization Prospects, Tech. rep., United Nations, New York, NY; 2014.

[48] Martínez-Zarzoso I, Maruotti A. The impact of urbanization on CO2 emissions: evidence from developing countries. Ecol Econ 2011;70(7):1344–53. https://doi.org/10.1016/j.ecolecon.2011.02.009.

[49] Zhou B, Thies S, Gudipudi R, Lüdeke MKB, Kropp JP, Rybski D. A Gini approach to spatial CO2 emissions [submitted for publication].

[50] Chertow MR. The IPAT equation and its variants. J Ind Ecol 2000;4(4):13–29. https://doi.org/10.1162/10881980052541927.

[51] Dietz T, Rosa EA. Effects of population and affluence on CO2 emissions. Proc Nat Acad Sci USA 1997;94(1):175–9. https://doi.org/10.1073/pnas.94.1.175.