ABSTRACT

Most parts of Africa’s infrastructure are yet to be built. Where and how these new buildings are constructed matters since today’s decisions will last for decades. The resulting morphology of cities has lasting implications for a city’s metabolism and energy needs. Estimating and projecting these needs has always been challenging in Africa due to the lack of data. Yet, given the sweeping urbanisation expected in Africa over the next three decades, this obstacle needs to be overcome to guide cities towards a trajectory of sustainability and resiliency. Based on the location and surface area of nearly 200 million buildings on the continent, we estimate the inter-building distance of almost six thousand cities. Satellite imagery enables the construction of urban form indicators to compare African cities’ elongation, sprawl and emptiness. We establish the BASE model, where the mean distance between buildings is a functional relation to the number of Buildings and their average Area, as well as the Sprawl and the Elongation of its spatial arrangement. The mean distance between structures in cities -our proxy for its energy demands related to mobility- grows faster than the square root of its population, resulting from the combined impact of a sublinear growth in the number of buildings and a sublinear increase in building size and sprawl. We show that when a city doubles its population, it triples the energy demand related to commutes.

1 Introduction

The world’s future is in cities. Projections estimate almost 7 out of 10 people will live in a city by 2050. Whilst many parts of the world have already undergone urbanisation, the next three decades will bring sweeping changes in African cities. An additional 950 million will become urbanites by 2050 in Africa. More people will need more buildings, future homes, schools, hospitals, markets, and all the other daily life stops. Where and how these buildings are constructed matters since today’s decisions will last for decades. The resulting morphology of cities -in other words, whether it is sprawling or compact, monocentric or polycentric, fragmented or contingent- has lasting implications for a city’s metabolism and its energy needs [1,2,3,4].

So far, efforts to tackle these questions have been waylaid by the lack of data. Isolated analyses of sprawl or fragmentation using population density data of major African cities exist, but smaller and intermediary cities are often excluded. A troubling phenomenon since most of the future urbanisation in the forthcoming decades will arise in these very cities. Much of the collective knowledge of urban form in African cities and future energy needs are based on samples of, at best, 100 cities.

To address these data gaps, we capitalise on the newly available Google AI Africa Open Buildings dataset [5], which maps every building on the continent. The granularity of this data, combined with prior work that maps over 5600 urban agglomerations and our mathematical models of urban form - enables us to estimate African cities’ metabolism with a never before seen accuracy.
2 Results

Here, we construct a set of indicators to characterise urban morphology based on the geographical distribution and size of millions of buildings in African cities [6]. The objective is to relate the morphology of cities to distance indicators (e.g., sprawl, elongation, polycentricity), and thus, to the urban “metabolism” [7, 8, 9]. In total, we use the coordinates and surface of 183 million buildings that have been constructed in nearly 6,000 cities in Africa (Appendix A). For each city identified by Africapolis [10], we measure the mean inter-building distance, which is our proxy for commuting energy consumption (Appendix A). Based on the expression for the average distance between two points inside a circle, a functional equation for the mean distance $D_i$ between two buildings inside city $i$ is given by

$$D_i = \frac{128}{45\pi} \sqrt{B_i A_i S_i E_i},$$

where $B_i$ is the number of buildings, and $A_i$ is their average area. $S_i$ is the sprawl index, which is the space between buildings, where smaller values come from a compact city. $E_i$ is the elongation of a city (or anisometry [8]), where smaller values indicate a round shape and higher values suggest a more elongated urban polygon (Appendix B). The sprawl $S_i$ and elongation $E_i$ are comparable across cities of different sizes, based on the maximum distance between buildings (Figure 1).

Figure 1: A - Cities may be elongated, sprawling, or both, depending on the location of their buildings. B - Footprint of buildings observed across nine cities with a similar population. Cities with an elongated shape have a larger diameter meaning that the mean distance between buildings is longer. Cities with more sprawl tend to have more spacing between buildings, contributing to longer distances. C - Observed increase in the number of buildings $B_i$, their area $A_i$, the sprawl $S_i$ and the elongation $E_i$ as the population of the city increases. D - Cumulative fraction of the number of buildings (top) and surface (bottom) of a city formed by buildings bigger than some area.

The BASE model decomposes distances in cities into four multiplicative components, where it is possible to analyse them separately [11]. The model gives an estimate for distances between buildings if a city is round and compact, $D_\star$, obtained in Equation [1] with $E_i = S_i = 1$. This way, $D_\star$ reflects how distances increase due to the city’s footprint (number and size of buildings). We define the fragmentation of a city as $\psi_i = D_i / D_\star = \sqrt{S_i E_i} \geq 1$, the ratio between the observed and minimum distances, in other words, the extent to which a city departs from a packed pattern (e.g., the space between buildings green space, blue space, roads). More fragmented cities have longer distances due to factors other than more or bigger buildings. We use cities’ elongation, sprawl, and fragmentation to detect what contributes to longer distances and characterise cities’ morphology across Africa. The physical terrain and the political boundaries play a crucial role in how cities grow into elongated or sprawled patterns. Urban areas with greater altitude variation tend to be more elongated and have greater sprawl (Appendix C). Cities near an international border are growing at a faster speed than other cities, and they also tend to be more elongated [12]. Polycentrism is related to the functionality of cities [13,14] and its morphology. More polycentric cities tend to have elongated shapes, greater sprawl and longer distances (Appendix D).

The number of buildings within a city grows with population, with roughly one extra building every 2.6 people. However, that number decreases slightly with size. A city with ten times the population has 9.6 times the number of buildings (Appendix E), so the number of buildings in a city grows sublinearly [15,16,17]. The buildings’ footprint
Measuring the morphology of African cities

comprises 10,000 constructed square kilometres in the continent (put together, it is roughly the surface of Lebanon). The majority are small residential buildings (68.5% of the buildings are less than 50 square metres, and only 0.3% have an area bigger than 500 square metres). However, big buildings, although small in quantity, might represent a large part of the constructed surface of a city (Appendix F). Buildings in Africa have an average surface of 55 square metres. However, the mean area of buildings also varies with city size. Roughly 18% of the constructed surface of Africa is buildings larger than 250 square metres. In Abidjan, for example, only 5% of the buildings are bigger than 250 square metres, but they total 30% of the constructed surface of the city (Figure 1). For example, a city with ten times the population has buildings that are, on average, 17% bigger, the result of a disproportionate presence of big buildings in larger cities (Appendix F). The same applies across the whole continent. In West Africa, for instance, a city with ten times the population has buildings that are 40% bigger. Large cities in the US and OECD countries are denser [13], but this is only observed in North Africa. In the rest of African regions, the footprint of a city increases superlinearly since cities have fewer buildings per person, but those buildings tend to be bigger. In West Africa, for instance, a city with ten times the population has 12.5 times the footprint (Appendix E).

Many cities are elongated, especially urban areas dominated by small buildings, close to physical barriers, international borders or more polycentric (Figure [1]). However, as the population increases, distances grow and become critical, so cities experience intense competition for space and make better use of it [19]. The mean distance between buildings is, on average, 2.8 times larger because of the non-circular shape of African cities. However, in West, South and Central Africa, larger cities are more compact (smaller sprawl S) and tend to be less elongated (smaller E), so even if they occupy more space, they settle it in a more efficient manner (Appendix E).

Distances in large cities grow sublinearly with population size. Without a scaling impact on the size of buildings and on elongation and sprawl, the mean distance between buildings should grow with the square root of the population (which would be observed if personal space remains constant for small and large cities). However, by fitting equation $D_i = \alpha D P_i^{\beta D}$ results range between $\hat{\beta}_C = 0.400$ for cities in Central Africa to $\hat{\beta}_N = 0.574$ for cities in North Africa. In North Africa, people occupy more space at an individual level in large cities. That extra space results from fewer but larger buildings and more elongated and sprawled cities. The same is not true across the rest of the continent. In West Africa, for example, larger cities are less elongated and have a smaller sprawl. It slightly counters the growth of distances due to a higher number of larger buildings in large metropolitan areas. Cities face two opposing forces as they grow. People demand more infrastructure, so more and larger buildings. But also, cities tend to be rounder and more compact to reduce distances and efficiently use space (Appendix E).

It is possible to characterise cities by observing only their centre or the densest location (Figure 2). Looking only at the city centre shows that large cities tend to have more than 30 or 40% of the surface built-up, as opposed to small cities, where less than 10% of its centre has been constructed (Appendix G). Thus, the sprawl near the centre in a city with ten times more people is 63% smaller.

The centre of a city might be “empty”, meaning that even when it is the most-dense spot in the city, the footprint might be pretty small, suggesting that the whole urban area has very low density. The emptiness of a city centre increases its sprawl and the metabolism of cities (Appendix G), which may result from a city having only a limited number of (mostly small) constructions. But city size is also related to the structure of a city by its centre. Buildings near the centre tend to be bigger in a larger city (Figure 2). For example, a city with ten times more population has 22% bigger buildings near the centre.

Some African countries, such as Niger or Chad, will double their 2020 population before 2050 [20]. Due to the urbanisation process and population growth, African cities will keep growing at an unprecedented speed, and some might reach a population of 80 or even 100 million inhabitants [21]. With a population of 80 million inhabitants, a city could have 18 million buildings (roughly two-thirds of the number of buildings there are today in Nigeria) with a footprint of 175 square kilometres and an average distance between buildings above 85 km (Appendix H). The burden on city dwellers of such an urbanisation process could force the city to grow in a roughly circular, compact and more vertical shape, urbanising green and blue areas. When a city grows, more people suffer longer distances. Thus, a city with ten times the population requires 34 times more energy in transport (Appendix H). By the time African cities double their 2020 population, the average distance between buildings will increase 45%, but more people will experience longer distances, so the energy consumed in transport in those cities will be 2.9 times the current levels (Appendix H). Thus, some African cities rely on reducing the number of daily journeys and spatial mismatch and on fast mass transport to avoid collapsing for excessive congestion and energy demands.

3 Discussion

Novel data based on the footprint of each building in Africa has given us the power to quantify the elongation and sprawl of a city and measure the mean distance between buildings -our proxy for the energy consumed in transport. Our
Measuring the morphology of African cities

Figure 2: A - The centre of a city, identified as its most dense location, and its 1 km vicinity is formed by some buildings that may be bigger than the rest of the city (forming a pyramid city) or even smaller (forming a valley), and the vicinity might have a small or large footprint. It is possible to characterise cities by looking only at the city centre. B, C, D - More populous cities have more and bigger buildings by their centre, so the footprint near the centre is also bigger. Among cities of more than one million inhabitants, at least 30% of its centre is built. E - Most cities (except for some cities mostly in East Africa) have less sprawl with lower levels of emptiness. Thus, efficient use of space near the city centre is crucial for reducing city sprawl.

The demographic transition and urbanisation process in Africa will be an unprecedented dynamic. Most of Africa’s infrastructure is yet to be built. Today, a strategic policy for urban expansion will set the trajectory of cities on the right track for given energy consumption, emissions, pollution, congestion, land use, and its ability to cope with climate change, such as extreme weather events like droughts, floods and heatwaves. Although small cities often lack urban expansion strategies, it is in large cities where the burden of sprawl and elongation is more significant. In the top 1% largest cities in Africa is where 80% of the total energy consumed in intercity transportation occurs. Furthermore, only 9% of African cities are coastal, but one-third of Africa’s top 1% cities are on a coast, where fragmentation increases. A 2050 plan for population, land use and mass transport in large cities is now an urgent element for constructing a sustainable continent [22].

Promoting compact land use in cities with bigger buildings (both in terms of area and height) arranged in a compact manner near the centre and preserving and fostering the accessibility of green open spaces are vital policies to improve cities’ liveability, although the urban heat island intensity might become more relevant and pervasive [8]. The novel data and the BASE model to measure indicators of the morphology of cities enabled us to classify cities based on their shape and detect groups of cities where the average distance between buildings (therefore, the energy required for transportation) is much longer than needed.

4 Methods

4.1 Data construction

Defining metropolitan areas is quite challenging and often depends on considerations and parameters [23, 24, 25]. Africapolis applies the same definition for a metropolitan area at a continental level, enabling it to analyse and compare cities between different countries [10]. Buildings’ locations were extracted from the recently launched Google Open Buildings dataset, https://sites.research.google/open-buildings/. The buildings’ footprints were obtained using a deep learning model with high-resolution satellite imagery (50 cm pixel size) [26]. We extract the buildings’ footprints that are located within Africapolis polygons and keep the attributes of the building’s centre location (latitude and longitude), the confidence score, and the building footprint area. We assign the unique identifier of the Africapolis urban agglomeration it belongs to (Appendix A). This way, we obtained the location and attributes of the buildings’ footprint of 6,849 African urban agglomerations (Figure 3). We computed street network metrics [27] and terrain
We define the sprawl of a city as everything else that increases distances in cities besides the number of buildings, their size, and their arrangement. Imagine that instead of buildings of a city, we have toy bricks of different sizes. There are infinitely many configurations to arrange those bricks keeping the maximum distance fixed (so keeping the diameter fixed), meaning that many brick configurations lead to the same maximum distance $D_i$, and this is the "diameter" of the city, that is, the longest distance between any two buildings and $A_i$ is the "minor axis" of a city (Appendix B). By considering $B_i$ and $A_i$ as the smallest and the largest radius, we obtain a scale-free coefficient $E_i$ that captures how elongated is the shape of the city (Figure 1). Cities mainly grow from the bottom up, and adjust to the topography, barriers and road infrastructure, so in general, the shape of cities is not circular [19]. Our data enables us to measure the mean distance $D_i$, the number of buildings $B_i$, and their size $A_i$. We construct $S_i$ and $E_i$ for each city. With values of $E_i$ close to one, the city’s shape is nearly circular, and higher values mean more elongated shapes. With small values of $S_i$, the city has a small sprawl, so buildings are arranged compactly. Thus, with small $E_i$ and $S_i$, the city has roughly a circular shape and compacted buildings. The impact of $E_i$ and $S_i$ is to increase distances, taking a tight circle as the basis. We construct our BASE model based on the formula for the expected distance between two points inside a circle. The average distance between any two points inside a circle with area $a$ is given by $128\sqrt{\pi/(45\pi)}$, so distances grow proportional to the radius. Equation (1) is inspired by the average distance between points inside a circle, where instead of the radius, we use the city’s footprint, $B_iA_i$, and two shape parameters with a multiplicative impact, $S_iE_i$.

Inspired by an ellipse, we measure the elongation of a city. Thus, with $E_i$, we capture how elliptical the city’s shape is. The mean distance between two points inside an ellipse has no closed solution [29], but an approximation can be constructed by considering the ratio between the major and the minor axis. Thus, we define the elongation $E_i$ as

$$E_i = \frac{\sqrt{\pi M_i}}{2\sqrt{B_i A_i}}$$  \hspace{1cm} (2)

where $M_i$ is the “diameter” of the city, that is, the longest distance between any two buildings and $2\sqrt{B_i A_i}/\pi$ is the smallest possible diameter of a circle with $B_iA_i$ as a footprint. Smaller values of $E_i$ mean that the ratio between the smallest and the largest radius is similar, so the city’s shape is more circular. Larger values mean more elongated shapes. In equation (2), $M_i$ is the “major axis”, and $2\sqrt{B_i A_i}/\pi$ is the “minor axis” of a city (Appendix B). By considering $E_i$ to be the ratio between two distances, we obtain a scale-free coefficient $E_i \geq 1$ concerning the number or area of buildings. That means that if a city is a scaled version of another, for example, with four times the number of buildings (or buildings four times the size), then distances also grow, including doubling the maximum distance and doubling the smallest diameter, so $E_i$ remains the same.

We define the sprawl of a city as everything else that increases distances in cities besides the number of buildings, their area and elongation. From equation (2) we get that

$$S_i = \frac{45^2 \pi^{3/2} D_i^2}{2^{13} M_i \sqrt{B_i A_i}} = \gamma \frac{D_i^2}{M_i \sqrt{B_i A_i}}$$  \hspace{1cm} (3)

for $\gamma = 45^2 \pi^{3/2}/2^{13} \approx 1.38$.

Imagine that instead of buildings of a city, we have toy bricks of different sizes. There are infinitely many configurations to arrange those bricks keeping the maximum distance fixed (so keeping the diameter fixed), meaning that many brick configurations lead to the same maximum distance $D_i$.
configurations have the same elongation. In one extreme, the bricks “fill” the area with the corresponding major axis, but in the other extreme, the largest distance is observed only between a few bricks. Therefore, the sprawl $S_i$ captures those possible brick arrangements.

Our technique decomposes the mean distance between buildings into four multiplicative components. Two components ($B_i$ and $A_i$) are measured directly from the data, and we have constructed a mathematical expression for $S_i$ and $E_i$, the sprawl and elongation. Thus, we have an exact expression that equates the mean distance between buildings in a city with four urban indicators (Equation 4).

$$D_i = \frac{128}{45\pi} \left( \frac{B_i A_i}{S_i E_i} \right)^{1/2}. \quad (4)$$

One way to interpret the elongation and the sprawl of a city is that the mean distance between buildings in cities increases proportionally to $\sqrt{E_i}$, so if a city has an elongation value of $E_i = 4$, then the mean distance between buildings is double because the city is elongated. Similarly for the sprawl $S_i$.

### 4.3 Minimum distances and fragmentation

The number of buildings in a city is the most obvious reason why distance grows in larger cities. More people means more buildings which then translates into larger distances. The same occurs with building size, so we “discount” those two reasons. Equation 4 enables us to detect how small distances could be if the city was compact and had a circular shape, with area $B_i A_i$ and with $E_i = 1$ and $S_i = 1$. The smallest mean distance $D_i^*$ is given by

$$D_i^* = \frac{128}{45\pi} \sqrt{B_i A_i}. \quad (5)$$

We construct the fragmentation index $\psi_i$ of a city, by

$$\psi_i = \frac{D_i}{D_i^*} = \sqrt{S_i E_i}, \quad (6)$$

that indicates how distances in city $i$ are longer due to an elongated shape and more spacing between buildings. Fragmentation is a coefficient $\psi_i \geq 1$ that is also scale independent to the number and area of buildings.

### 4.4 Measuring polycentrism

Cities are increasingly characterised by polycentricity, that is, the presence of multiple interconnected centres [30]. Measuring polycentrism is based on three methodological stages: delineating urban regions, identifying subcentres and then applying some mathematical function to obtain an index for a city [30]. Different techniques have been used based on infrastructure data, such as satellite images, the density of points of interest, road network or buildings data [31, 32, 33, 34] and also based on mobility data or travel surveys [35, 36]. Measuring polycentrism depends not only on the type of data and the criteria applied for defining centres and subcentres but also on the possibly divergent methods of measuring polycentricity [14].

Here, big remote sensing data is used to identify the spatial clustering of buildings in cities. The kernel density estimates the intensity of the number of buildings and the constructed surface per unit area [37]. For example, the technique has been used for constructing a continuous surface representing the intensity of human activity based on points of interest in a city [38]. The kernel density is a cumulative function obtained by adding a decaying surface for each building. Formally, for some point $x$ in space, we define the kernel as

$$k(x) = \sum_{j=1}^{n} w_j f(d_{x,j}), \quad (7)$$

where $n$ is the number of buildings, $w_j$ are the weights of each building, taken to be its area, and $d_{x,j}$ is the distance between the point $x$ and the centroid of the building, and $f$ is a decreasing function, known as the smoothing kernel, here taken to be a Gaussian function. The result is a surface over each urban area that highlights parts with more or larger constructions (peaks) considered centres of the city [38].

Based on the kernel density, a contour adjacent relation tree is constructed [39], where each node on the tree is a new contour. Contour trees represent spatial relationships between the contours of the kernel surface and summarise
relationships at different levels [40]. Separate peaks (urban centres) are identified as branches on a tree, which are connected depending on the different contour levels of the surface. The procedure gives $N$ branches, where $N = 1$ is a monocentric city. Each branch has three indicators: height (corresponding to the kernel estimate), area (representing the total surface of the city that belongs to that branch) and volume (obtained by multiplying the area and height of each branch).

Let $Br_i$ be the number of branches of city $i$ and let $v_k$ be the volume of each branch in decreasing order (so that $v_1 \geq v_2 \geq \cdots \geq v_{Br_i}$). The polycentrism index $\phi_i$ is defined as

$$\phi_i = \frac{1}{v_1} \sum_{k=1}^{Br_i} k v_k.$$  \hfill (8)

The index $k$ inside the sum of equation 8 helps increase the value of $\phi_i$ with the number of branches. By dividing by $v_1$ in equation 8, a comparable index across cities is obtained (Appendix D). If $\phi_i = 1$ then $i$ is monocentric. A city with two distant and equal-sized centres (so that they belong to different branches) has $\phi_i = 3$ (with three equal-sized centres $\psi_i = 6$, and so on). The procedure also gives the volume tree of a city, where the dimension of each branch (horizontal axis) and its height (vertical axis) represent the centres of the city [41] [39].

Larger cities tend to be more polycentric, formed by more branches with more separation between them and a larger relative volume. However, medium-size cities may also be highly polycentric, reflecting fragmented urban areas.

### 4.5 Impact of city size

The infrastructure and the socio-economic outputs of a city vary according to many factors, and city size has been detected to be a critical aspect of cities [42] [15] [16]. We characterise the four city indicators of the BASE model and detect if they vary according to city size. We fit equations like $B_i = \alpha P_i^\beta$ between the number of buildings $B_i$ and the population $P_i$ for some coefficients $\alpha$ and $\beta$, and similarly for the other indicators. Values of $\beta \approx 0$ indicate that city size has little or no impact on the corresponding indicator. Values of $\beta \approx 1$ indicate a linear growth, and values below (and above) $\beta = 1$ are a sublinear (or superlinear) impact of city size (Appendix E).

The scaling coefficients show a sublinear relation with the number of buildings $B_i$ (so fewer buildings per person as cities grow), sublinear with the area of buildings $A_i$ (so bigger buildings in larger cities). Also, larger cities tend to have slightly more sprawl, and the elongation is statistically the same across cities of different size (Table 1).

| variable                  | $\alpha$ | $\beta$ |
|---------------------------|----------|---------|
| $B_i$ - number of buildings| 0.423    | 0.029   |
| $A_i$ - area of buildings | 24.85    | 1.29    |
| $S_i$ - sprawl             | 0.725    | 0.067   |
| $E_i$ - elongation         | 5.925    | 0.337   |
| $D_i$ - mean distance      | 6.126    | 0.479   |

Table 1: Observed scaling coefficients at the continental level.
where the population has a scaling coefficient \( s \beta_D = 0.440 \). Suppose larger cities were just a scaled version of a small city. In that case, the number of buildings per person should be constant (so \( \beta_D = 1 \)), and the rest of the scaling coefficients should be zero, reflecting the same building size, sprawl and elongation, so that the mean distance should grow with the square root of the population. However, cities are not a scaled version of a small city, and most components vary with city size.

Beyond distances, it is possible to decompose the footprint and the fragmentation of a city. The footprint \( \theta_i \) can be decomposed as \( \theta_i = B_i A_i = \alpha_\theta \rho^{\beta_B + \beta_A} \), so for example, the footprint is superlinear in West Africa but sublinear in North Africa. The fragmentation \( \psi_i \) can be decomposed as \( \sqrt{S_i E_i} = \alpha_\psi \rho^{(\beta_S + \beta_E)/2} \), where \( \alpha_\psi = \sqrt{\alpha_S \alpha_E} \) and the population has a scaling coefficient \( (\beta_S + \beta_E)/2 \). Thus, larger cities in North Africa are more fragmented than smaller cities mainly due to a higher sprawl, but the opposite happens in Central Africa, where the sprawl and elongation is smaller in larger cities, so they tend to be less fragmented (Appendix E).

### 4.6 Characterising cities based on the city centre

Considering the distribution of the urban footprint at a distance \( R = 1 \) km from its centre enables us to observe the sprawl across cities based on the same shape. Comparing the same shape across different urban polygoms means we can ignore the elongation and focus only on the sprawl. The sprawl of a city inside a circle, referred to as the emptiness \( T_i^{(R)} \geq 1 \), is defined as the ratio between the surface of the city centre and the part that is constructed, given by

\[
T_i^{(R)} = \frac{\pi R^2}{B_i^{(R)} A_i^{(R)}},
\]

where \( B_i^{(R)} \) and \( A_i^{(R)} \) are the number and average area of buildings inside that circle. The emptiness \( T_i^{(R)} \) corresponds to the observed sprawl inside a circle, and a higher emptiness means that a smaller surface of the centre has been constructed (Appendix G). Although longer distances than \( R = 1 \) km may also be considered for observing cities at their centre, we observe that a circle with a 1-km radius is the largest circle that fits almost always inside the polygon of all cities. Bigger circles often are not fully contained inside the polygon of cities, so analysing the footprint depends on some elongation, whereas smaller circles have a smaller building sample in cities.

### Table 2: Observed scaling coefficients across regions in Africa

| \( B_i \)   | North  | West   | East   | Central | South  |
|------------|--------|--------|--------|---------|--------|
|            | 0.926***| 0.957** | 1.042** | 0.990***| 1.005***|
|            | (0.013) | (0.011) | (0.010) | (0.017) | (0.014) |
| \( A_i \)  | 0.044***| 0.144** | 0.035** | 0.076***| 0.018  |
|            | (0.009) | (0.010) | (0.007) | (0.012) | (0.010) |
| \( S_i \)  | 0.106***| -0.026* | 0.026*  | -0.150***| -0.078***|
|            | (0.012) | (0.009) | (0.013) | (0.014) | (0.019) |
| \( E_i \)  | 0.071***| -0.017* | -0.001  | -0.109***| -0.065***|
|            | (0.010) | (0.007) | (0.010) | (0.015) | (0.019) |

* \( p < 0.001; ** p < 0.01; * p < 0.05 \)
4.7 Data construction

We use open geospatial data sets for this analysis: urban boundaries, buildings’ locations, street network data, and elevation data. The boundaries of African urban agglomerations were downloaded from the Africapolis website in November 2021 [10]. The downloaded dataset contains 516M building footprints, across an area with sprawl. Africapolis applies the same definition for a metropolitan area, enabling us to compare cities at a continental level. According to the Africapolis website, an urban agglomeration unit combines satellite and aerial imagery and official demographic data such as censuses and other cartographic sources. The delineation of an urban agglomeration boundary is based on a spatial approach that combines physical limits, the continuity of the built-up area, and demographic criteria to provide a standardised continental spatial definition of urban areas. All urban agglomerations considered have more than 10,000 inhabitants. Also, since our procedure combines different geographical data, other filters are applied to the cities:

- Drop 1,202 cites with less than 2000 buildings (not enough data in the city)
- Drop 123 cities with elongation $E_i > 20$ and seven cities with sprawl $S_i > 4$, reflecting that polygons and buildings do not spatially align. The mean elongation is $\mu_E = 7.5$ with $\sigma_E = 4.5$, and the mean sprawl is $\mu_S = 0.94$ with $\sigma_S = 0.83$, so cities with a higher elongation or sprawl are outliers.
- Drop 24 cities with a footprint per person above 60 square metres. The mean footprint per person is $\mu_\theta = 19$ square metres per person, with $\sigma_\theta = 20$.

Buildings’ locations were extracted from the recently launched Google Open Buildings dataset, https://sites.research.google/open-buildings/, which contains 516M building footprints, across an area of 19.4M km², 64% of the African continent [5]. The buildings’ footprints were obtained using a deep learning model with high-resolution satellite imagery (50 cm pixel size) [26]. This dataset is packed into 136 different files in comma-separated values format. It contains information on the location of the building’s footprint centre in geographic coordinates (latitude and longitude), its area in square meters, a Plus Code corresponding to the centre of the building, and a confidence score that informs about the confidence level of the building detection, and the geometry of the building’s footprint. See https://maps.google.com/pluscodes/ for more information on Plus Codes.

We downloaded the 136 Open Buildings files (49.8 GB) and implemented a geoprocessing workflow in R software [43] to extract the buildings’ footprints within Africapolis polygons. We keep the attributes of the building’s centre location (latitude and longitude), the confidence score, and the building footprint area. We assign the unique identifier (agglosID) of the Africapolis urban agglomeration it belongs. In this way, we obtained the location and attributes of the buildings’ footprints of 8,489 African urban agglomerations. There are 871 Africapolis urban agglomerations without Google Open Building data. We’ve noticed that Morocco, Western Sahara, Libya, Chad, Cameroon, and South Sudan, as well as North Kivu (Democratic Republic of Congo), Cabo Delgado Province (Mozambique), and the Southwestern region in Sudan, all lack buildings’ footprints in the Google Open Buildings dataset.

We computed street network metrics (the average street length and the total number of nodes or intersections) using the Python library OSMnx [27] with the Africapolis urban agglomeration polygons. We used that information to compute the number of buildings per street node for each urban agglomeration. Finally, we also computed some terrain metrics, such as the difference in elevation between the highest and lowest point, the average slope, and the average height within each Africapolis urban agglomeration polygon using the NASADEM digital elevation model in Google Earth Engine [28].

4.8 Measuring the elongation and sprawl of a city

Based on the eccentricity of an ellipse, we construct a metric using the ratio between the “major” and “minor” axis of a city. The major axis of a city is $M_i$, the maximum observed distance between buildings. If all city buildings were arranged next to the other in a circular shape, they would form a circle with area $B_i A_i$ and with radius $\sqrt{B_i A_i/\pi}$, considered to be the minor axis of a city. There is no closed formula to measure the mean distance between two points inside an ellipse [29]. A good approximation is given by $128\theta/(45\pi)$, where $\theta = \sqrt{a/b}$, and where $a$ is the major axis and $b$ is the minor axis of the ellipse, so $\theta$ captures the discrepancy between $a$ and $b$. The formula is similar to the distance between two points inside a circle of radius $b$ but multiplied by a factor $\theta \geq 1$ that captures the shape of the ellipse. The formula is exact for $\theta = 1$ and works well for small values of $\theta$. Therefore, in an ellipse, distances grow roughly proportional to $\theta$. For city $i$, we construct its elongation $E_i$ as the ratio between the maximum observed distance between buildings and the minimum possible radius of a city. As with an ellipse, values of $E_i$ close to one mean a more circular shape, whereas higher values mean elongated cities across some axis. Also, distances in the city are expected to grow proportional to $\sqrt{E_i}$, so the expression aligns with the BASE model.

4.9 Terrain, elevation and physical factors

The growth of cities is often bounded by physical barriers, such as mountains or coasts. Thus, when we observe that a city is very elongated, it is frequently because of rugged terrain and differences in elevation. To measure the impact of physical barriers and rough terrain on the elongation of a city, we measure the distance to the nearest coast, the
Measuring the morphology of African cities

Figure 4: The (modified) eccentricity of an ellipse $e'$ is defined as the ratio between the semi-major and semi-minor axis, both squared. Similarly, we define the city eccentricity $\theta_i$ as the ratio between the maximum observed distance (taken as the major axis) and the minimum radius of a circle with $B_i$ buildings with an average surface $A_i$.

Figure 5: Number of branches (top) and polycentrism (bottom) observed in a city according to its population (horizontal axis).

Point with the lowest and the highest altitude in the city, and differences in altitude. This set of variables enables us to correlate the city metrics and detect whether being close to a coast increases the urban form indicators. Cities with high elevation or differences in altitude suggest proximity to mountains and the presence of physical barriers and rugged terrain.

4.10 Polycentrism in cities

Cities frequently develop and grow in a polycentric manner, that is, by merging different urban areas or adding new centres. Here, we measure polycentrism using the spatial distribution of buildings, the corresponding kernel density surface and the contour tree it forms. Results show that the number of branches of the corresponding contour tree increase with city size (Figure 5). However, those new branches add a small relative volume for small cities. In general, only cities above 100 thousand inhabitants are polycentric, where those extra branches add volume separate to the city’s centre.

There is a wide variation in the levels of polycentrism in cities. Take, for example, cities with roughly one million inhabitants. The levels of polycentrism vary from 1 to values above 20, suggesting that it is not only size that affects the polycentric growth of cities.

4.11 Impact of city size on urban form

People from larger cities tend to produce more patents, have a higher income per capita, and suffer more crime and some types of diseases [15]. People from larger cities migrate less and are more likely to return after moving [44]. Smaller cities usually require more infrastructure per person, such as the road surface or the petrol stations [15]. A mathematical expression to capture the impact of city size is obtained by adjusting the equation

$$Y_i = \alpha P_i^\beta,$$

where $Y_i$ is the variable of interest for city $i$. Then, $Y_i$ could be the number of restaurants in city $i$. We obtain the coefficients $\alpha$ and $\beta$ through some technique (for example, via a regression or minimising the sum of square errors). With $\hat{\beta} > 1$ the results are called “superlinear” and indicate that large cities have higher values of the variable $Y$ per capita (since the per capita rate $Y_i / P_i$ is given by $\hat{\alpha} P_i^{\hat{\beta} - 1}$).

With $\hat{\beta} < 1$ results are “sublinear” and with $\hat{\beta} \approx 1$ city size has little or no impact on the per capita rate of that city. Therefore, the coefficient of interest is usually $\beta$ and values above or below $\beta = 1$ are critical. It has been observed that for some social indicators, $\beta = 1.15$ is frequently obtained, which means that when comparing two cities, $i$ and $j$, where the population of $j$ is ten times the population of $i$, then, the expected values of $Y_j$ are approximately $10^{1.15} = 14.13$ times larger, and, on a per capita basis, city $j$ has 1.413 more of $Y$ than city $i$. Thus, a city ten times larger has 14.13 times more...
4.12 The size of buildings across Africa

Most buildings across Africa have a tiny footprint. A small footprint corresponds to low-rise constructions, and only in buildings with a large footprint, there is uncertainty as to whether it is a high or a low-rise building. Therefore, the distribution of the footprint of buildings is also an indicator of infrastructure, including construction materials, overcrowding, and resilient structures, among many economic indicators. We see that most buildings, particularly from smaller cities, are quite small. For example, most buildings in Bangui are smaller than 25 square metres (something like a square with 5 metres on each side). We compare the distribution of building size between cities. However, outliers in terms of building size are relevant, and they do alter the city indicators. For example, consider two cities with 25,000 buildings (roughly 100,000 inhabitants) each, with a similar abundance of small buildings. Observing that one of those two cities has 250 large buildings, so 1% of them changes the way we look at that city since it could correspond to state capital, for instance. Also, results vary if by “large” buildings we mean with a footprint of more than 100 square metres or more than 500 square metres. Therefore, we consider the percentage of the built-up area constructed from buildings larger than some threshold \( \kappa > 0 \). We vary the values of \( \kappa \) from 20 metres (tiny buildings) to 5000 metres (so, large buildings).

The vast majority of African buildings are small. There are, however, substantial differences across cities. Large buildings in some cities are scarce, particularly in small cities, while large buildings are frequently observed in some cities such as Johannesburg, Cape Town, Abidjan or Cairo.

To detect the impact of city size in the number of large and small buildings, we consider a threshold area \( \rho \) and measure, for city \( i \), the number of buildings smaller than \( \rho \), say \( s_i(\rho) \), and the number of buildings larger than \( \rho \), say \( l_i(\rho) \). Then, we fit the Poisson regression

\[
s_i(\rho) = \alpha_s P_i^{\beta_s}(\rho), \quad l_i(\rho) = \alpha_l P_i^{\beta_l}(\rho).
\]

For example, values of \( \beta_s(\rho) \) below 1 indicate an abundance of buildings smaller than \( \rho \) in smaller cities, whereas values above 1 indicate an abundance of small buildings in larger cities. And the same of \( \beta_l(\rho) \) for buildings larger than \( \rho \). Different values of \( \rho \) result in a different division of small or large buildings.

Results show that the number of buildings in a city is sublinear (so smaller cities tend to have more buildings per person than a larger city), but the relationship changes with building size (Figure 6 top). Considering buildings with more than 100 square metres, the scaling coefficient is approximately one, meaning the same number of buildings per person in small and large cities. As buildings grow, they become scarce in small cities, so the scaling coefficient increases. For buildings larger than 500 square metres, the scaling coefficient is \( \beta_{500} = 1.12 \), meaning that for a city that is ten times the size, the number of buildings
larger than 500 square metres is $10^{1.12} = 13.2$ times larger. Thus, larger cities have bigger buildings.

The footprint of a city constructed from buildings larger than some area (top) for different areas (horizontal axis). Scaling coefficient $\beta$ (vertical) of the area in the city constructed based on buildings larger than some area (top) for different areas (horizontal axis).

The presence of large buildings is related to city size, but also, the location of those large buildings within a city plays a role. We distinguish some cities where the centre has a high frequency of large buildings, and the area decays as the distance to the city centre increases, thus, forming a “pyramid” type of city. In contrast, a different kind of city is observed when large buildings are spread across the urban polygon.

We construct a metric that captures whether a city has a pyramid or a more flat dispersion of buildings. But we also want to consider the relative size of buildings. That is, a city might be regarded as more pyramid-like if the largest buildings in the city are by its centre, even if those buildings are relatively small. For measuring a relative factor of how the city has a pyramidal distribution, we construct an index based on the relative size of buildings in cities. The coefficient is constructed by the ratio between the average size of buildings within the city centre and the average size of buildings in that city. Formally, let $\mu_i$ be the average size of buildings in city $i$, and let $\nu_i$ be the average size of buildings within a radius of 1 km within the city centre. Then, the flatness coefficient $\Delta_i$ is given by

$$\Delta_i = \frac{\nu_i}{\mu_i}$$

where $\Delta_i \approx 1$ means that buildings in a city have a similar footprint in the city centre. Larger values of $\Delta_i$ suggest a pyramid-type of city, and smaller values of $\Delta_i$ suggest that the city centre has smaller buildings than the outskirts.

The city centre is identified as the point with the highest weighted density of buildings. That is, we consider a kernel density estimate of the point process formed by the coordinates of buildings, weighted by their footprint. Thus, the density surface gives the constructed area within some radius at each $x, y$ coordinate. We then identify the city centre as the location with the highest built density. Although the city centre could be defined differently (for example, by analysing the centroid of the polygon, looking for some central business district or looking at the building closer to others), our technique identifies the location in the city with the highest level of constructed surface.

By comparing the average size of buildings within the centre to the rest of the city, we obtain an index that identifies the dispersion of large buildings within a city. If the recognised centre has smaller buildings, it corresponds to a city centre with small but very compact buildings. If buildings within the city centre are of equal size to the rest of the city, it means a small variability in the size of buildings, corresponding to a flat city. We consider the average size of buildings within a radius of 1 km within the city centre. The flatness coefficient $\Delta_i$, combined with other aspects such as city size, mean building size and emptiness, distinguishes the internal structure of cities.

4.14 Distance and energy consumption

The mean distance between buildings in a city is our proxy for the energy that the city requires for transportation. Given the average distance between buildings on city $i$, expressed as a function of the population by $D_i = \alpha_T P_i^{\beta_D}$, we estimate that the total energy consumed for transport in the city is proportional to $T_i = \alpha_T P_i^{\beta_D+1}$ for some $\alpha_T > 0$. This way, $P_i$ experiences a commuting distance in the city that is proportional to $D_i^{\beta_D}$. Although we cannot estimate values of $\alpha_T$ based on our data, we observe that at a continental level, $\beta_D = 0.532$. Therefore, we estimate that the total energy consumed for transport in African cities grows superlinearly, with a coefficient of 1.532. Thus, a city with ten times the population requires 34 times more energy in transport because the city is larger, so with more and bigger buildings, distances grow.

The BASE model enables us to predict the number of buildings and the size of a city with some population $P_d$ and also estimates a city’s sprawl and elongation index. Most countries in Africa will double their 2020 population, some
a city that doubles its population will require 2.97 times more energy, but a city in Central Africa will only require 2.63 times more energy since cities have less sprawl and become less elongated as they increase in size. The opposite happens in North Africa, where larger cities tend to have higher levels of fragmentation.

even before 2050, and some cities might reach a population of 80 or even 100 million inhabitants [21]. Due to the continent’s demographic expansion and urbanisation process, most cities will more than double their population in the decades to come. And when a city doubles in size, the total energy requirements in the continent will increase three times. Our estimate varies by region. In North Africa,
Table 3: Model number of buildings $B$

$$ y = \log B $$

| Term               | Coefficient | Std. Error | $p$-value |
|--------------------|-------------|------------|-----------|
| (Intercept)        | 2.4331      | 0.0979     | ***       |
| log(Polycentrism)  | 0.4115      | 0.0150     | ***       |
| RegionEast         | 0.0821      | 0.0182     | ***       |
| RegionNorth        | 0.0655      | 0.0211     | ***       |
| RegionSouth        | 0.5718      | 0.0230     | ***       |
| log(dif elev m)    | 0.1140      | 0.0057     | ***       |

** $p < 0.001$; * $p < 0.01$; * $p < 0.05$

Table 4: Model building area $A$

$$ y = \log A $$

| Term               | Coefficient | Std. Error | $p$-value |
|--------------------|-------------|------------|-----------|
| (Intercept)        | -10.372     | 0.026      | ***       |
| log(dif elev m)    | 0.008       | 0.002      | ***       |
| log(BuildingsPerNode) | -0.005   | 0.002      | ***       |
| log(Polycentrism)  | -0.004      | 0.004      | ***       |
| a100               | 0.877       | 0.016      | ***       |
| TotalFootprintCentre1km | 0.000  | 0.000      | ***       |
| MeanArea1000m      | 0.010       | 0.000      | ***       |
| IsPyramid          | -0.370      | 0.006      | ***       |
| RegionEast         | 0.001       | 0.005      | ***       |
| RegionNorth        | -0.058      | 0.006      | ***       |
| RegionSouth        | -0.032      | 0.006      | ***       |
| RegionWest         | -0.071      | 0.005      | ***       |
| log(street_length_avg) | 0.009    | 0.004      | *          |
| log(Population_2015) | -0.008    | 0.002      | ***       |

** $p < 0.001$; * $p < 0.01$; * $p < 0.05$

Table 5: Model sprawl $S$

$$ y = \log S $$

| Term               | Coefficient | Std. Error | $p$-value |
|--------------------|-------------|------------|-----------|
| (Intercept)        | -2.510      | 0.099      | ***       |
| RegionEast         | 0.013       | 0.017      |           |
| RegionNorth        | 0.064       | 0.020      | **        |
| RegionSouth        | 0.091       | 0.022      | ***       |
| RegionWest         | -0.034      | 0.018      |           |
| MeanBuildingArea   | 1521.843    | 1073.992   |           |
| log(dif elev m)    | 0.049       | 0.006      | ***       |
| log(nodes)         | 0.114       | 0.006      | ***       |
| a100               | 0.039       | 0.063      |           |
| IsPyramid          | 0.223       | 0.033      | ***       |
| TotalFootprintCentre1km | -0.000   | 0.000      | ***       |
| MeanArea1000m      | -0.002      | 0.001      | ***       |
| log(Polycentrism)  | 0.399       | 0.015      | ***       |
| log(street_length_avg) | 0.295    | 0.013      | ***       |
| ICoast             | 0.071       | 0.014      | ***       |
| IBorder             | 0.058       | 0.017      | ***       |
| log(Population_2015) | 0.025    | 0.010      | ***       |

** $p < 0.001$; * $p < 0.01$; * $p < 0.05$

Table 6: Model elongation $E$

$$ y = \log E $$

| Term               | Coefficient | Std. Error | $p$-value |
|--------------------|-------------|------------|-----------|
| (Intercept)        | -0.211      | 0.088      | *          |
| RegionEast         | 0.064       | 0.015      | ***       |
| RegionNorth        | -0.072      | 0.018      | ***       |
| RegionSouth        | -0.189      | 0.020      | ***       |
| RegionWest         | -0.109      | 0.016      | ***       |
| MeanBuildingArea   | 907.518     | 949.954    |           |
| log(dif elev m)    | 0.078       | 0.005      | ***       |
| log(nodes)         | 0.092       | 0.005      | ***       |
| a100               | 0.071       | 0.056      |           |
| IsPyramid          | 0.085       | 0.029      | ***       |
| TotalFootprintCentre1km | -0.000   | 0.000      | ***       |
| MeanArea1000m      | -0.002      | 0.001      | *          |
| log(Polycentrism)  | 0.218       | 0.013      | ***       |
| log(street_length_avg) | 0.275    | 0.012      | ***       |
| ICoast             | 0.050       | 0.012      | ***       |
| IBorder             | 0.037       | 0.015      | ***       |
| log(Population_2015) | 0.022    | 0.008      | ***       |

** $p < 0.001$; * $p < 0.01$; * $p < 0.05$
### Measuring the morphology of African cities

\[ y = \log \psi \]

| Term                      | Coefficient | Standard Error |
|---------------------------|-------------|----------------|
| (Intercept)               | -1.360      | 0.081          |
| RegionEast                | 0.039       | 0.014          |
| RegionNorth               | -0.004      | 0.017          |
| RegionSouth               | -0.049      | 0.018          |
| RegionWest                | -0.072      | 0.015          |
| MeanBuildingArea          | 1214.681    | (878.774)      |
| log(dif_elev_m)           | 0.063       | (0.005)        |
| log(nodes)                | 0.103       | (0.005)        |
| a100                      | 0.055       | (0.052)        |
| IsPyramid                 | 0.154       | (0.027)        |
| TotalFootprintCentre1km   | -0.000      | (0.000)        |
| MeanArea1000m             | -0.002      | (0.001)        |
| log(Polycentrism)         | 0.308       | (0.012)        |
| log(street_length_avg)    | 0.285       | (0.011)        |
| ICoast                    | 0.061       | (0.011)        |
| IBorder                   | 0.048       | (0.014)        |
| log(Population_2015)      | 0.023       | (0.008)        |

**R\(^2\)**: 0.622  
**Adj. R\(^2\)**: 0.620  
**Num. obs.**: 5625

***p < 0.001; **p < 0.01; *p < 0.05***  

**Table 7: Model fragmentation \( \psi \)**

\[ y = \log \Delta \]

| Term                      | Coefficient | Standard Error |
|---------------------------|-------------|----------------|
| (Intercept)               | -0.6059     | (0.0392)       |
| RegionEast                | 0.0364      | (0.0097)       |
| RegionNorth               | -0.0301     | (0.0097)       |
| RegionSouth               | 0.0649      | (0.0128)       |
| RegionWest                | -0.0816     | (0.0097)       |
| TotalFootprintCentre1km   | -0.0000     | (0.0000)       |
| FootprintPerPerson         | 0.0032      | (0.0003)       |
| MeanBuildingArea          | -1780.1001  | (131.6482)     |
| log(Population_2015)      | 0.0763      | (0.0039)       |

**R\(^2\)**: 0.2331  
**Adj. R\(^2\)**: 0.2320  
**Num. obs.**: 5625

***p < 0.001; **p < 0.01; *p < 0.05***  

**Table 9: Ratio building size inside city centre, \( \Delta \)**

\[ y = a100 \]

| Term                      | Coefficient | Standard Error |
|---------------------------|-------------|----------------|
| (Intercept)               | -0.1562     | (0.0209)       |
| RegionEast                | 0.0472      | (0.0036)       |
| RegionNorth               | 0.1099      | (0.0041)       |
| RegionSouth               | 0.0935      | (0.0045)       |
| RegionWest                | 0.0780      | (0.0037)       |
| MeanBuildingArea          | 7678.7422   | (201.9923)     |
| log(dif_elev_m)           | -0.0036     | (0.0012)       |
| log(nodes)                | 0.0131      | (0.0013)       |
| IsPyramid                 | 0.0010      | (0.0070)       |
| TotalFootprintCentre1km   | 0.0000      | (0.0000)       |
| MeanArea1000m             | 0.0007      | (0.0002)       |
| log(Polycentrism)         | 0.0052      | (0.0031)       |
| log(street_length_avg)    | 0.0071      | (0.0028)       |
| ICoast                    | 0.0130      | (0.0029)       |
| IBorder                   | 0.0091      | (0.0037)       |
| log(Population_2015)      | -0.0027     | (0.0020)       |

**R\(^2\)**: 0.8979  
**Adj. R\(^2\)**: 0.8976  
**Num. obs.**: 5625

***p < 0.001; **p < 0.01; *p < 0.05***  

**Table 10: area from buildings larger than 100m**
Measuring the morphology of African cities

|                          | $y = \log \theta_{1km}$ | $y = \log \theta_{2km}$ |
|--------------------------|--------------------------|--------------------------|
| (Intercept)              | 6.4191***                | 3.5406***                |
|                          | (0.0864)                 | (0.0859)                 |
| log(Polycentrism)        | -0.6074***               | -0.6226***               |
|                          | (0.0178)                 | (0.0177)                 |
| MeanBuildingArea         | -7007.4782***            | -5026.2724***            |
|                          | (1415.9295)              | (1406.6931)              |
| a100                     | 0.6553***                | 1.2729***                |
|                          | (0.0826)                 | (0.0821)                 |
| log(dif_elev_m)          | 0.0486***                | 0.1111***                |
|                          | (0.0071)                 | (0.0070)                 |
| RegionEast               | 0.2033***                | 0.1593***                |
|                          | (0.0229)                 | (0.0228)                 |
| RegionNorth              | 0.2968***                | 0.1778***                |
|                          | (0.0259)                 | (0.0258)                 |
| RegionSouth              | 0.2934***                | 0.5418***                |
|                          | (0.0289)                 | (0.0287)                 |
| RegionWest               | 0.2516***                | 0.1469***                |
|                          | (0.0241)                 | (0.0239)                 |
| ICoast                   | 0.0135                   | 0.0729***                |
|                          | (0.0184)                 | (0.0183)                 |
| MeanArea1000m            | 0.0008***                | 0.0013                   |
|                          | (0.0012)                 | (0.0012)                 |
| Capital                  | -0.0937                  | -0.3469***               |
|                          | (0.0635)                 | (0.0631)                 |
| IsPyramid                | -0.3640***               | -0.1067*                 |
|                          | (0.0432)                 | (0.0429)                 |
| log(Population_2015)     | 0.5560***                | 0.8415***                |
|                          | (0.0074)                 | (0.0074)                 |

$R^2$                     0.6233  0.8130
Adj. $R^2$                0.6224  0.8126
Num. obs.                5509  5509

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 11: Ratio building size inside city centre
Measuring the morphology of African cities

### References

[1] Xuemei Bai, Alyson Surveyer, Thomas Elmqvist, Franz W Gatzweiler, Burak Güneralp, Susan Parnell, Anne-Helene Prieur-Richard, Paul Shrivastava, Jose Gabriel Siri, Mark Stafford-Smith, et al. Defining and advancing a systems approach for sustainable cities. *Current Opinion in Environmental Sustainability*, 23:69–78, 2016.

[2] Shlomo Angel, Sara Arango Franco, Yang Liu, and Alejandro M Blei. The shape compactness of urban footprints. *Progress in Planning*, 139:100429, 2020.

[3] Susan Parnell, Thomas Elmqvist, Timon McPhearson, Harini Nagendra, and Sverker Sörlin. Introduction: Situating knowledge and action for an urban planet. *The urban planet: knowledge towards sustainable cities*, pages 1–16, 2018.

[4] Erneson A Oliveira, Vasco Furtado, José S Andrade, and Hernán A Makse. A worldwide model for boundaries of urban settlements. *Royal Society Open Science*, 5(5):180468, 2018.

[5] Wojciech Sirko, Sergii Kashubin, Marvin Ritter, Abigail Annkah, Yasser Salah Eddine Bouchareb, Yann Dauphin, Daniel Keysers, Maxim Neumann, Moustapha Cisse, and John Quinn. Continental-scale building detection from high resolution satellite imagery. *arXiv preprint arXiv:2107.12283*, 2021.

[6] Reid Ewing and Shima Hamidi. Measuring urban sprawl and validating sprawl measures. *Washington, DC: National Institutes of Health and Smart Growth America*, 2014.

[7] Marc Barthelemy. The statistical physics of cities. *Nature Reviews Physics*, 1(6):406–415, 2019.

[8] Bin Zhou, Diego Rybski, and Jürgen P Kropp. The role of city size and urban form in the surface urban heat island. *Scientific Reports*, 7(1):1–9, 2017.

[9] Yunfei Li, Sebastian Schubert, Jürgen P Kropp, and Diego Rybski. On the influence of density and morphology on the Urban Heat Island intensity. *Nature Communications*, 11(1):1–9, 2020.
Measuring the morphology of African cities

Table 14: Scaling coefficients across regions in Africa
Measuring the morphology of African cities

[10] François Moriconi-Ebrard, Dominique Harre, and Philipp Heinrigs. Urbanisation dynamics in West Africa 1950–2010, 2015.

[11] Ramana Gudipudi, Diego Rybski, Matthias KB Lüdeke, Bin Zhou, Zhu Liu, and Jürgen P Kropp. The efficient, the intensive, and the productive: Insights from urban Kaya scaling. Applied Energy, 236:155–162, 2019.

[12] OECD. Population and morphology of border cities. 1(21), 2019.

[13] Chen Zhong, Markus Schlüsper, Stefan Müller Arisona, Michael Batty, Carlo Ratti, and Gerhard Schmitt. Revealing centrality in the spatial structure of cities from human activity patterns. Urban Studies, 1, 2015.

[14] Bartosz Bartosiewicz and Szymon Marcińczak. Investigating polycentric urban regions: Different measures–different results. Cities, 105:102855, 2020.

[15] Luís MA Bettencourt, Jose Lobo, Dirk Helbing, Christian Kuhnert, and Geoffrey B West. Growth, innovation, scaling, and the pace of life in cities. PNAS, 104(17):7301–7306, 2007.

[16] Markus Schlüpfer, Luís MA Bettencourt, Sébastian Graswin, Mathias Raschke, Rob Claxton, Zbigniew Smoreda, Geoffrey B West, and Carlo Ratti. The scaling of human interactions with city size. Journal of the Royal Society Interface, 11(98):20130789, 2014.

[17] Luís MA Bettencourt. The origins of scaling in cities. Science, 340(6139):1438–1441, 2013.

[18] Rémi Louf and Marc Barthelemy. How congestion shapes cities: from mobility patterns to scaling. Scientific Reports, 4(1):1–9, 2014.

[19] Michael Batty. The size, scale, and shape of cities. Science, 319(5864):769–771, 2008.

[20] Rafael Prieto Curiel, Carmen Cabrera-Arnau, and Steven R Bishop. Scaling beyond cities. Frontiers in Physics, 10, 2022.

[21] Daniel Hoornweg and Kevin Pope. Population predictions for the world’s largest cities in the 21st century. Environment and Urbanization, 29(1):195–216, 2017.

[22] Reid Ewing. Is Los Angeles-style sprawl desirable? Journal of the American Planning association, 63(1):107–126, 1997.

[23] Hernán D Rozenfeld, Diego Rybski, José S Andrade, Michael Batty, H Eugene Stanley, and Hernán A Makse. Laws of population growth. Proceedings of the National Academy of Sciences, 105(48):18702–18707, 2008.

[24] Clémentine Cottineau, Erez Hatna, Elsa Arcaute, and Michael Batty. Diverse cities or the systematic paradox of urban scaling laws. Computers, Environment and Urban Systems, 63:80–94, 2017.

[25] Elsa Arcaute, Erez Hatna, Peter Ferguson, Hyejin Youn, Anders Johansson, and Michael Batty. Constructing cities, deconstructing scaling laws. Journal of the Royal Society Interface, 12(102):20140745, 2015.

[26] Wojciech Sirko, Sergii Kashubin, Marvin Ritter, Abigail Annkah, Yasser Salah Eddine Bouchareb, Yann Dauphin, Daniel Keysers, Maxim Neumann, Moustapha Cisse, and John Quinn. Continental-scale building detection from high resolution satellite imagery, 2021.

[27] Geoff Boeing. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems, 65:126–139, 2017.

[28] NASA JPL. NASADEM Merged DEM Global 1 arc second V001 [Data set],. 2020.

[29] Michelle Parry and Ephraim Fischbach. Probability distribution of distance in a uniform ellipsoid: Theory and applications to physics. Journal of Mathematical Physics, 41(4):2417–2433, 2000.

[30] Ryan Thomas, Stephan Schmidt, and Stefan Siedentop. Toward comparative polycentricity scores: Assessing variations in regional delineation and subcenter identification. Environment and Planning B: Urban Analytics and City Science, 1, 2021.

[31] Zhiwei Xie, Xinyue Ye, Zihao Zheng, Dong Li, Lishuang Sun, Ruren Li, and Samuel Benya. Modeling polycentric urbanization using multisource big geospatial data. Remote Sensing, 11(3):310, 2019.

[32] Prosper Issahaku Korah, Tony Matthews, and Deanna Tomerini. Characterising spatial and temporal patterns of urban evolution in Sub-Saharan Africa: The case of Accra, Ghana. Land Use Policy, 87:104049, 2019.

[33] Yue Deng, Jiping Liu, Yang Liu, and An Luo. Detecting urban polycentric structure from POI data. ISPRS International Journal of Geo-Information, 8(6):283, 2019.

[34] Hannes Taubenböck, Ines Standfuß, Michael Wurm, Angelika Krehl, and Stefan Siedentop. Measuring morphological polycentricity—a comparative analysis of urban mass concentrations using remote sensing data. Computers, Environment and Urban Systems, 64:42–56, 2017.
Measuring the morphology of African cities

[35] Narisra Limtanakool, Martin Dijst, and Tim Schwanen. A theoretical framework and methodology for characterising national urban systems on the basis of flows of people: empirical evidence for France and Germany. *Urban Studies*, 44(11):2123–2145, 2007.

[36] Jameson L Toole, Michael Ulm, Marta C González, and Dietmar Bauer. Inferring land use from mobile phone activity. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, pages 1–8, 2012.

[37] Adrian Baddeley, Ege Rubak, and Rolf Turner. *Spatial point patterns: methodology and applications with R*. CRC press, 2015.

[38] Peter Diggle. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press, Florida, US., third edition edition, 2014.

[39] Chengming Li, Peipei Guo, Pengda Wu, and Xiaoli Liu. Extraction of terrain feature lines from elevation contours using a directed adjacent relation tree. *ISPRS International Journal of Geo-Information*, 7(5):163, 2018.

[40] Eric Guilbert. Multi-level representation of terrain features on a contour map. *Geoinformatica*, 17(2):301–324, 2013.

[41] Jussi Sakari Klemelä. *Smoothing of multivariate data: density estimation and visualization*, volume 737. John Wiley & Sons, 2009.

[42] Denise Pumain and Marianne Guerois. Scaling laws in urban systems. In *Santa Fe Institute, Working Papers*, pages 1–26, 2004.

[43] R Core Team. R: A language and environment for statistical computing, 2018.

[44] Rafael Prieto Curiel, Luca Pappalardo, Lorenzo Gabrielli, and Steven Richard Bishop. Gravity and scaling laws of city to city migration. *PloS one*, 13(7):e0199892, 2018.